ERIC REPORT RESUME

ED 010 270 2-28-67 24 (REV)

NONOVERT REINFORCED CLOZE PROCEDURE.

BLOOMER, RICHARD H. * AND OTHERS
DDK11050 UNIVERSITY OF CONNECTICUT, STORRS

CRP-2245
BR-5-0534
- -66
EDRS PRICE MF-$0.18 HC-$5.04 152P.

*CLOZE PROCEDURE, REINFORCEMENT, *READING COMPREHENSION, *LANGUAGE ABILITY, MEASUREMENT TECHNIQUES, TESTING PROGRAMS, STORRS, CONNECTICUT

THE STUDY MEASURED THE EFFECTS OF TRAINING IN A SERIES OF CLOZE PROCEDURE EXERCISES UPON READING COMPREHENSION. STUDENTS WERE INSTRUCTED TO FILL IN BLANK SPACES IN WRITTEN TEXT. THIS ACT OF CLOSURE ITSELF WAS CONSIDERED TO BE A NONOVERT REINFORCEMENT. STUDENTS (1,411) IN 49 CLASSROOMS PARTICIPATED IN THE STUDY. A CROSS-SECTIONAL GROUP WAS STUDIED FOR 3 WEEKS, AND A LONGITUDINAL GROUP FOR 12 WEEKS. THE STUDENTS WERE FROM GRADES 5, 7, 9, AND 11. THE AUTHORS CONCLUDED THAT READING COMPREHENSION WAS NOT IMPROVED BUT SUGGESTED OTHER POSSIBLE USES OF THE PROCEDURE. (JK)

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

NON-OVERT REINFORCED CLOZE PROCEDURE

A Project Report

of

Office Of Education 2245

ED010270

Staff

Richard H. Bloomer, Project Director

Vincent Louthan, Research Associate

Andrew J. Heitzman, Consultant

Reed Mitchell, University Research Assistant

University of Connecticut

1966

TABLE OF CONTENTS

Chapter		Page
I	Introduction	1
II	Literature	10
III	Procedure	15
IV	Pilot Study I	34
V	Pilot Study II	43
VI	The Effects of Non-Overt Reinforced Cloze Procedure Upon Comprehension by Grade and Sex . .	48
VII	The Effects of Non-Reinforced Cloze Procedure Upon the Ability to Make Closures	59
VIII	The Effect of Non-Overt Reinforced Cloze Procedure Upon Reading Comprehension	77
IX	The Relation of Non-Reinforced Cloze Procedure To the Intelligence of the Respondent	87
X	Areas for Future Exploration of Cloze Procedure .	100
	BIBLIOGRAPHY	123
	LIST OF TABLES	iii

LIST OF TABLES

Table		Page
I	STUDIES OF CLOZE READABILITY	10a
II	STUDIES OF CLOZE AND GENERAL READING ACHIEVEMENT .	11a
III	STUDIES OF CLOZE AND SPECIFIC READING COMPREHENSION (SAME MATERIALS)	11c
IV	STUDIES OF CLOZE AND VERBAL ABILITY AND IQ	12a
V	CLOZE AS A TEACHING DEVICE	12b
VI	CROSS-SECTIONAL EXPERIMENT SHOWING NUMBER OF PUPILS BY GRADE LEVEL AND SEX	16
VII	MEAN CORRECT COMPREHENSION QUESTIONS BY BLOCK AND GRADE	29
VIII	ANALYSIS OF VARIANCE OF BLOCKS BY GRADES	29
IX	MEANS AND 't' RATIOS OF COMPREHENSION QUESTIONS CORRECT.	37
X	MEANS AND 't's OF CLOZE BLANKS CORRECT (PERCENTAGE - ALL STUDENTS)	38
XI	MEAN NUMBER OF CORRECT COMPREHENSION QUESTIONS BY TREATMENT AND DELETION TYPE	47
XII	ANALYSIS OF VARIANCE OF MEAN COMPREHENSION QUESTIONS CORRECT BY TREATMENT AND DELETION TYPE .	47
XIII	MEAN CORRECT RESPONSES TO COMPREHENSION QUESTIONS BY GRADE LEVELS AND TYPE OF RESPONSES FOR ALL SUBJECTS	51a
XIV	RECALL COMPREHENSION TEST SCORES BY DELETION TYPE	53a
XV	MEAN CORRECT RESPONSES TO COMPREHENSION QUESTIONS BY GRADE LEVELS AND TYPES OF DELETIONS	
	FOR BOY SUBJECTS.	55a
	FOR GIRL SUBJECTS	55b
XVI	MEAN NUMBER OF CORRECT CLOSURES FOR EXERCISES 1, 10 and 22 BY DELETION TYPE	65

Table Page

XVII 't' TEST COMPARISON OF NUMBER OF CORRECT CLOSURES IN EXERCISE 1 WITH EXERCISES 10 and 22 .. 65

XVIII SIGNIFICANT 't' RATIOS BETWEEN NUMBER OF CORRECT CLOSURES FOR EXERCISE 1 BY DELETION TYPE 67a

XIX SIGNIFICANT 't' RATIOS BETWEEN NUMBER OF CORRECT CLOSURES FOR EXERCISE 10 BY DELETION TYPE 67a

XX SIGNIFICANT 't' RATIOS BETWEEN NUMBER OF CORRECT CLOSURES FOR EXERCISE 22 BY DELETION TYPE 67a

XXI MEAN FINAL CLOZE SCORES FOR TEST DELETION BY STUDENTS COMPLETING EACH TYPE OF EXERCISE DELETION 67b

XXII 't' TEST COMPARISONS OF CONTROL GROUP (ZERO DELETIONS) WITH TYPE OF EXERCISE DELETION BY TEST DELETION TYPE 68a

XXIII PRE-TEST DIFFERENTIAL APTITUDE TEST MEAN SCORES--D.A.T. SUBTEST 79a

XXIV POST TEST. THE IOWA TEST OF BASIC SKILLS MEAN SCORES 79b

XXV MEAN NUMBER OF CORRECT COMPREHENSION QUESTIONS BY EXERCISE AND DELETION TYPE 80a

XXVI SIGNIFICANT 't' TEST COMPARISONS OF CORRECT COMPREHENSION QUESTIONS FOR EXPERIMENTAL AND CONTROL GROUP (TYPE 0) BY EXERCISE 81

XXVII CORRELATIONS OF READING TEST SCORES WITH NUMBER OF CORRECT CLOSURES FOR EXERCISES 1, 10 and 22 BY DELETION TYPES 82

XXVIII CORRELATION OF READING TEST SCORES WITH NUMBER OF CORRECT COMPREHENSION ITEMS CORRECT FOR EXERCISES 1, 10 and 22 83

XXIX CALIFORNIA TEST OF MENTAL MATURITY MEAN SCORES BY DELETION TYPE 87a

XXX MEAN CORRELATIONS FOR CALIFORNIA TEST OF MENTAL MATURITY SCORES WITH NUMBER OF COMPREHENSION QUESTIONS CORRECT BY TYPE OF SYNTACTIC DELETION 87a

XXXI CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CORRECT COMPREHENSION QUESTIONS FOR EXERCISES 1, 10 and 22 91

Table Page

XXXII MEAN CORRELATION OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF COMPREHENSION QUESTIONS CORRECT ON RECALL FINAL BY TYPE OF SYNTACTIC DELETION 93

XXXIII CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CORRECT COMPREHENSION QUESTIONS ON RECALL FINAL FOR EXAMINATIONS 1, 10 and 22 93

XXXIV MEAN CORRELATIONS FOR CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CORRECT CLOSURES IN TRAINING EXERCISES BY TYPE OF SYNTAX 95

XXXV MEAN CORRELATION FOR CALIFORNIA TEST OF MENTAL MATURITY WITH TOTAL CORRECT CLOSURES ON THE FINAL CLOZE TEST BY TYPE OF SYNTACTIC DELETION . . 95

XXXVI CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CORRECT CLOSURES FOR EXERCISES 1, 10 and 22 97

XXXVII CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CLOSURES BY SYNTACTIC TYPE ON CLOZE FINAL TEST 98

CHAPTER I

INTRODUCTION

The cloze procedure has in recent years been receiving considerable attention. This attention has primarily been focused upon two aspects of the use of cloze procedure, namely, readability and reading comprehension testing. The third aspect, that of using the cloze procedure as a technique for teaching reading *per se* in various reading improvement situations, has also begun to be explored. The present study deals with this aspect of the cloze procedure as a teaching device, but by virtue of the cloze procedure itself, the other aspects will in some measure also be investigated.

Definition of the Cloze Procedure. To define the cloze procedure in terms of its construction is a relatively simple matter. To determine what this particular structure does when applied to pupils is considerably more complex. The cloze procedure is basically a systematically mutilated passage delivered to students with the instructions to fill in the blanks with the word which is missing from the text. The systematic deletion is generally accomplished by selecting one of the first words in the passage by a random procedure and then following this selection with the deletion of every *Nth* word. Commonly every fifth word is deleted from cloze procedure material.

For the present study, for purposes to be outlined below, every tenth word is deleted. In the original study of cloze by Taylor it was shown that within these limits it made relatively little difference whether every fifth or every tenth word was deleted for readability purposes. In addition to the systematic deletion, there is typically a minimum deletion sample size required to give some reliability to the measure. Particularly is this true when one is concerned with cloze procedure as a readability or comprehension testing device. The minimum which seems to be commonly accepted is 50 deletions or 250 words. There is not an apparent maximum to the length of a cloze procedure passage, but certainly such factors as fatigue and extinction suggest that this may well be an area to explore.

There are essentially two ways to handle cloze procedure materials. The first is as a testing device; the second as a teaching device. The first involves no reinforcement, or knowledge of the correct response on the part of the responder. In essence, the individual fills in the cloze blanks and turns in his paper, and may or may not see it again. But if he does, it is simply in the shape of a mark or a score.

The second is through a technique of reinforcement; here the individual achieves a knowledge of the correct response which he has made. This knowledge of correct response allows the individual then to accomplish.

One must consider, however, that non-reinforced cloze procedure is not completely without reinforcement. Certainly the individual who fills in a cloze blank can, by his own means by matching his language pattern thus produced by his closure, reinforce himself. Hence non-reinforced cloze procedure is in effect non-overt reinforced cloze procedure. It is the effect of this non-reinforced cloze procedure which is the topic of this present study.

In essence, the concern will be to attempt to measure the effects of a prolonged series of non-overtly-reinforced cloze procedure exercises upon reading comprehension and upon the ability to make closures.

Readability. The initial use of the cloze procedure by Taylor[38] was to determine the readability of passages. The number of correct responses to the mutilation in a particular passage gives an indication of the difficulty of the passage. If such a passage is delivered to a group of people, and there is a significant difference between the mean number of correct closures of that passage and another passage delivered to the same or similar people, we can assume that the words which were not mutilated in the more difficult passage did not give a sufficient amount of information to the reader to allow him to make the closures without making errors. Essentially then, at one word in five deletions, we are saying that eighty percent of the words in one passage contained less information regarding the total meaning of the passage than in another case. We may also infer that either redundancy is reduced or that the language pattern used in the passage is unusual and does not allow the individual to guess the proper word using his available language ability. In either case the passage is more difficult to read.

Two important questions arise in relation to measuring readability by cloze. First, does it make any difference whether one scores synonyms as correct or whether one simply scores as correct the exact word deleted? Taylor's early work shows there is essentially no difference in the rank order of passages, regardless of the scoring system used.

The second question which arises is: Does the deleted word make any difference in the outcome of the readability estimates? Here Bormuth's study[3] which systematically deleted each one of the words in five different forms of a cloze procedure selection indicates there is no difference as a function of which words are left out if the deletions are done in the systematic cloze procedure fashion. Cloze procedure readability essentially is comparative, and what is needed in this particular area is some variety of standardization so that given a particular level of responding, *x* number of errors, or *x* proportion of errors, will represent a particular grade level or age level of reading for a particular child. Essentially the relationship between the proportion of deletion errors in a cloze procedure passage and the ability of the child at a particular level in school to read the material needs to be more fully explored.

Comprehension Testing. Just as different passages prove to be of differential difficulty for groups of people, so too does the same passage appear to be of differential difficulty for different people. With this as a premise, the cloze procedure can be used as a reading comprehension testing device. Large numbers of studies have used cloze procedure for testing reading comprehension. Early studies in this areas were by Taylor[42] and by Rankin.[28] With one or two notable exceptions, these studies find a reasonable correlation between the number of correct responses on a cloze procedure passage and reading comprehension test scores. The interesting thing about cloze procedure as a comprehension testing device is that it allows the analysis of the internal comprehension process as it is a technique for asking continuous comprehension questions during the reading of the passage; hence, in relatively long passages, may be analyzed for a differential in reading comprehension in various parts of the passage, or may be systematically altered to determine some of the specific variables which may enter into reading difficulty, both in relation to the nature of the material itself, and also in relation to the nature of changes in the individual reader as he goes through a particular passage. It is clear that reading comprehension depends upon the ability of an individual to predict words which have been deleted from a particular passage. Hence the aspect of reading comprehension measured is the concurrence of the language pattern of the reader with the language pattern of the writer. Where these two concur the reader is well able to estimate what words the writer will use and the amount of change which must take place in the reader in order to understand the particular passage is reduced. The cloze procedure seems to fit the concept of learning to comprehend by reading phrases and avoiding word by word silent reading, or even oral reading, of the passages. Where closures are readily made, the language ability of the individual is developed to the level of the particular writing. The concepts tend therefore to be fairly well formed and cloze procedure reading comprehension is a measure, not so much of how much the individual can learn from a passage but of how much he already knows about it when he enters the passage. The reverse is true when we consider the typical reading comprehension selection which consists of a paragraph followed by a series of questions about the paragraph which the individual must answer. Here we know relatively little about the process by which the individual comes up with a correct answer, and we are measuring not so much his reading comprehension ability directly but how much he is able to learn or retain from the particular passage. While these two aspects are quite obviously correlated and are even more highly correlated when a time limit is put on the normal reading comprehension situation, they appear also to be somewhat different--the one dealing with existing concepts, the other dealing with the ability of the individual to develop concepts from reading itself. Perhaps what is needed here is a definition of reading comprehension as either one or the other of these aspects. In essence, is reading comprehension to be measured by the number of extant concepts which an individual holds in common with the reading material, or is reading comprehension the ability of an individual to learn new concepts from the material? In either instance, the reading process via the cloze procedure where an individual must stop periodically and insert a word is a different procedure from the straight reading that we are accustomed to.

Cloze Teaching. It is this difference between the cloze procedure and the reading process which makes possible the use of cloze procedure as a teaching device for improving reading comprehension. For teaching purposes reinforced cloze procedure is essentially a continuous stream of short-answer questions, the answers to which must be derived from the reading material which surrounds the blanks with terminal reinforcement. The individual, therefore, who does a cloze procedure passage is required to focus somewhat more upon the material than the individual who is simply reading the material without regard for blanks. Therefore, if the concepts are already available in the learner, cloze procedure forces him to derive actively the concept from his own knowledge, whereas continuous reading allows the individual to absorb the concepts passively from the writer. The effects of the focusing required in cloze procedure are somewhat debatable. If an individual does not have the available language patterns and sufficient cues are not presented in the reading material, the individual may have difficulty understanding a cloze procedure passage; hence, there will be loss in comprehension. If the individual can manage to fill in most of the blanks in the cloze procedure material, the focusing on the material conceivably may have the effect of making the reader more attentive to the material he is reading. Such a procedure has the effect of reducing the rate at which individuals read but the overall time saved by good comprehension may more than make up for this time loss. On this basis several experiments have been performed which tend to lend some credence to this particular position.[1,32,15] Each of these studies involved the use of feedback to the pupils in terms of the correct answer which was delivered immediately after the pupil had completed the cloze exercise. The feedback of the correct response would give the individual reinforcement or corroboration for the responses he made and allow him to determine where his particular errors were made. For present purposes it is necessary to examine the cloze procedure somewhat more closely. If the cloze procedure is a series of short answer questions about a bit of material which is at once both the questions and the source for the answers, then, excepting format, the cloze procedure is very similar to programed instruction, the difference being that a large number of answers are made before the feedback occurs (rather than immediate feedback, as is the case with programed instruction). The Roosnick study[32] made use of this immediate feedback for each of the items and found that a very low error rate could be developed for materials designed for fifth and sixth grade pupils.

Let us for a moment examine the meanings of the word closure. Closure is a Gestalt term which suggests that when a situation is left unfinished the individual himself tends to finish it--in essence, to close the gaps, to make the material "good". Essentially this is an internal process and is not as such available to scientific scrutiny. However, if one conjectures on the meaning in psychological terms of "making things good", this goodness may be construed as a state of self-reinforcement. Closure and the subjectively described positive affect state which comes with getting an answer which to the responder fits the particular situation may well be a situation in which the individual can reinforce himself. In essence, if the language pattern which he develops fits both his known language pattern and makes sense to him in terms of what the writer has put down, he automatically gets an amount of feedback irrespective of whether objective feedback is available. If the preposition that closure represents self-reinforcement is in fact tenable, then any cloze procedure material would in turn be a teaching device, albeit perhaps not as effective as the same material where overt feedback is delivered.

Cloze Procedure and Linguistics. While Bormuth's study has shown that the systematic random deletion of words does not affect the scores one would obtain upon cloze procedure materials, an analysis by Bloomer and Heitzman of the errors made, by syntax of deletion, indicated a differential in the probability of error for the various syntactical elements. In effect this is saying that the deletion of words of a particular syntax over another makes a cloze passage more or less difficult. If we look at the deletions in terms of the contribution to the meaning of the passage we find that words of some particular syntaxes (in this case nouns and adjectives) tend to contribute more meaning to the passage than do other word forms, particularly the function words. The Bloomer and Heitzman study employed random deletions. By controlling the relative difficulty of the passage by the syntax of deletions, the ability of the individual to learn from the material and to match language patterns with the writer can be measured in relation to the contribution to meaning of words of varying syntaxes. In addition the nature of the substituted words (the error words) will indicate something of the nature of the language pattern of the individual.

The present study proposes to investigate an aspect of each of these four areas--cloze readability, cloze comprehension, cloze teaching, and cloze linguistics--by presenting twenty-four different selections, followed by comprehension questions, in one of the following cloze forms:

1. Control (no deletions)
2. Random deletions
3. Noun deletions
4. Verb deletions
5. Adjective deletions
6. Preposition and conjunction deletions
7. Article deletions

CHAPTER II

LITERATURE

The experimental literature dealing with the cloze procedure now approaches fifty articles and dissertations. An excellent review of these materials is given in Rankin's survey in the 1964 National Reading Conference Yearbook. The present review of the literature will restrict itself to studies pertinent to the present investigation. The treatment by Rankin is somewhat more exhaustive in the types of studies covered.

Readability. Table I gives the studies of cloze procedure as a technique for measuring readability. In reading the table, under the heading "Cloze Method", any word refers to the random deletion of words regardless of their syntax. Unless otherwise specified in this column, the cloze method is given either first or independently of the variable which is measured against it. Noun-verb deletions represent passages in which only nouns and verbs have been deleted. The evidence in Table I indicates that the cloze procedure can be used to measure readability. Cloze has been used with some success in several languages as well as English and when a passage is given several times a reliable estimate of reading difficulty which can be used for rank ordering is available. Essentially the work which has not yet been done on the cloze procedure as a readability technique is to establish the number of correct cloze responses to be anticipated at a particular grade level. In essence, standardization of the cloze readability technique is not available.

Table I

STUDIES OF CLOZE READABILITY

Variable	Cloze Method	Statistics	Author	Date
Dale-Chall and Flesch Formulae	Any Word Assorted passages including James Joyce and Gertrude Stein	Same Rank Easy by Readability Difficult by Cloze	Taylor	1953
Predicted Difficulty	3 Any Word, Korean	Same Rank Order	Taylor	1954
Judges Rating	8 Any Word, Japanese	rho - .83	Sukeyori	1957
Multiple Choice Comprehension Questions	9 Any Word Passages	rho - .92	Bormuth	1962

General Reading Achievement. The cloze procedure has been extensively examined as a technique for testing reading achievement. Table II gives a summary of the results of the various studies. The cloze procedure gives respectable correlations with reading achievement in each case save the Weaver and Kingston[46] study and of course cloze has been used as a reading comprehension testing device in the Stanford Achievement Test.

In the present study the cloze selections are compared against standardized reading tests given both pre-test and post-test. The fact that syntactical deletions of a single type are found in particular selections allows the determination of the relationship between reading comprehension scores and particular syntactical deletion, in addition to the any-word deletions which have been customarily used.

Cloze and Comprehension of the Same Materials. Table III summarizes a series of studies which relate the number of correct closures with comprehension tests covering the same materials. The relationships are relatively high except for the study by Friedman[15] which shows that when individual scores are related to comprehension test scores the correlations appear to be quite low. One would expect this as the scores of an individual are less stable than those of a group and less likely to represent a true score. At the same time it indicates that the reliability of cloze for individual testing may be relatively low.

Each of the cloze selections in the present study also has a comprehension test given immediately after it. As a consequence, relationships, both individual and group, can be determined for cloze material with deletions of varying syntaxes. In addition the relationship

Table II

STUDIES OF CLOZE AND GENERAL READING ACHIEVEMENT

Variable	Cloze Method	Statistic	Author	Date
Diagnostic Survey	Any Word Deletion		Rankin	1957
Story Comprehension		r - .29		
Vocabulary		r - .68		
Paragraph Comprehension		r - .60		
Cooperative Reading Test C 2	Any Word Deletion		Jenkinson	1957
Vocabulary		r - .78		
Level of Comprehension		r - .73		
Cooperative Reading Test C 2	Any Word Deletion		Fletcher	1959
Vocabulary		r - .63		
Level of Comprehension		r - .55		
Speed of Comprehension		r - .57		
Dvorak-Van Wagenen Diagnostic Examination of Silent Reading Ability	Any Word Deletion		Fletcher	1959
Rate of Comprehension		r - .59		
Diagnostic Reading	Post-Cloze Noun, Verb Deletions		Rankin	1962
Story Comprehension		r - .65		
Vocabulary		r - .45		
Paragraph Comprehension		r - .59		
Diagnostic Reading Test	Noun, Verb Deletions		Rankin	1963
Story Comprehension		r - .57		
Vocabulary		r - .42		
Paragraph Comprehension		r - .39		

(continued)

Table II (continued)

Variable	Cloze Method	Statistic	Author	Date
Diagnostic Reading Survey	Post-Cloze Deletion:		Greene	1964
Total Comprehension	Any Word	r - .51		
	Noun, Verb, Adjective, Adverb	r - .67		
Davis Reading Test	Two Cloze Tests:		Weaver and Kingston	1963
	Noun, Verb Deletions	r - .21		
		r - .36		
	Two Cloze Tests:			
	Any Word Deletion	r - .25		
		r - .51		
Stanford Achievement Test Paragraph Meaning	Any Word Deletion	r - .61-.74	Ruddell	1963
Metropolitan Achievement Test, Reading Section	Twenty Cloze Tests:		Friedman	1964
Vocabulary	Any Word Deletion	r - .63-.85		
Total Reading	Any Word Deletion	r - .71-.87		
Michigan Vocabulary Profile Test	Any Word Deletion	r - .56	Hafner	1963

Table III

STUDIES OF CLOZE AND SPECIFIC READING COMPREHENSION (SAME MATERIALS)

Variable	Cloze Method	Statistic	Author	Date
Comprehension Test	Any Word Deletion	r - .82	Jenkinson	1957
Comprehension Test	Post-Cloze: Any Word Deletion	r - .80	Taylor	1957
Comprehension Test	Post-Cloze: Noun, Verb Deletions	r - .78	Rankin	1957, 1959
Comprehension Test	Nine Tests, Any Word Deletions All Nine Above, Combined Score	r - .73-.84 r - .93	Bormuth	1962, 1963
Comprehension Test	Twenty Tests, Any Word Deletions Total Errors, Foreign Students Individual Scores	r - .90-.91 r - .24-.43	Friedman	1964
Pre-Reading Knowledge	Noun, Verb, Adverb Deletions	r - .92	Taylor	1957
Pre-Knowledge	Any Word Deletions, One Week Early	r - .86	Rankin	1957, 1959
Post-Knowledge Test	Following Pre-Test No Pre-Test	r - .25 r - .57	Bloomer and Heitzman	1962
Immediate Recall	Post-Cloze: Noun, Verb, Adverb Deletions	r - .80	Taylor	1957
Post-Reading Criterion Test	Post-Cloze: Noun, Verb Deletions	r - .78	Rankin	1957, 1959

between the various cloze forms and straight unmutilated reading of the same material in relation to the comprehension is determined.

Cloze Procedure and Personality. A study by Rankin has shown that cloze procedure scores are somewhat less reliable for introverts than for extroverts. In the present study the California Test of Personality is given to determine the interaction between personality variables and cloze deletions as well as reading comprehension with various personality characteristics.

Cloze Procedure, Verbal Ability and IQ. Table IV indicates that the cloze procedure is strongly related to the number of verbal measures as well as the number of intelligence tests or sub-tests. The measurement of IQ is included in the present study primarily as a support for these materials.

Cloze Procedure as a Teaching Device. Several studies have been done using the cloze procedure as an intermediary between pre-test and post-test, the purpose of the cloze procedure being primarily as an intervening variable. Each of the four studies reported indicates in one way or another that cloze procedure may be useful as a teaching device. In one instance there was no significant difference between the gain of a straight reading group and the gain in a cloze reading group; in the second instance significant differences were found in favor of people taking cloze procedure material as opposed to controls with no reading experience. In each of the studies reported, some kind of motivational device was used. It is very difficult to separate the motivation and the possible benefits accruing therefrom from the effect

Table IV

STUDIES OF CLOZE AND VERBAL ABILITY AND IQ

Variable	Cloze Method	Statistic	Author	Date
Air Forces Qualification Tests	Any Word Deletion	r - .72	Taylor	1957
Word Knowledge Sub-Test	Any Word Deletion	r - .85		
Arithmetic Reasoning	Any Word Deletion	r - .69		
IQ	Any Word Deletion	r - .69	Jenkinson	1957
A.C.E.			Fletcher	1959
Linguistic Sub-Test	Any Word Deletion	r - .72		
Quantitative	Any Word Deletion	r - .45		
Reasoning Factor	Any Word Deletion	r - .76	Weinfeld	1959
Verbal Factor		r - .70		
Theme Writing		r - .59		
Fluency of Expression		r - .54		
Word Fluency		r - .50		
Ideational Fluency		r - .40		
Otis IQ	Any Word Deletion	r - .73	Hafner	1963
Wechsler-Bellevue Intelligence Scale Information Sub-Test	Any Word Deletion	r - .56		
Thurstone's Understanding Communication	Any Word Deletions	r - .52	Greene	1964
	Noun, Verb Deletions	r - .61		

Table V

CLOZE AS A TEACHING DEVICE

Variable	Cloze Method	Statistic	Author	Date
Pre-Knowledge, Intervening Cloze	Post-Knowledge	$r - .37$	Bloomer and Heitzman	1964
Pre-Knowledge Test, Straight Reading	Pre-Knowledge Test, Any Word Deletion	No Significant Difference	"	"
Post-Knowledge Test	Post-Knowledge Test	No Significant Difference	"	"
Cloze Pre-Knowledge Test	No Pre-Knowledge Test Any Word Deletion	No Pre-Knowledge Superior	"	"
Post-Knowledge Test	Post Knowledge Test	$p \div .05$		
Diagnostic Reading Test	12 to 50 Cloze Any Word Deletion depending upon Quality of Response, vs. No Cloze Control		Bloomer	1962
Pre- and Post-Comprehension		Experimental Superior		
Pre- and Post-Reading Rate		No Significant Difference		
Pre- and Post-Vocabulary		No Significant Difference		
Predicted College Grades, Experimental and Control		Experimental Superior		
	200 Key Word Deletion Clozes	High Rate of Accurate Responses	Roossinck	1962
Control: 20 Mc-Call-Crabbs Standard Test Lessons	Experiment: 20 Cloze McCall-Crabbs Standard Test Lessons	Both Groups Gained No Significant Difference	Friedman	1964

of the cloze material itself. If we look upon cloze procedure as self-reinforcing we might anticipate that where motivation is reduced to a minimum, the motivation that may come from making correct cloze responses may produce similar effects. In the present study all attempts were made to reduce the motivation through the manipulation of the cloze procedure in one form or another. Aside from the mere passing out of papers in the classroom, the children were in no other way stimulated to do the cloze procedure in the present experiment. In addition, since they received no external feedback in relation to responses, motivation from rewards of being told they did well or poorly was eliminated from the experiment. This procedure allows us to determine whether the effects found in previous studies were a function of the combination of cloze procedure and motivation, or whether the results were more related to the cloze materials themselves. In addition, by using deletions of various syntaxes which in turn control the level of difficulty (Bloomer and Heitzman) we are better able to determine the effects of varying difficulties as well as the varying syntaxes in the present study.

SUMMARY

In general, the literature shows strong support for the use of cloze procedure as a readability tool. It is also clear from the studies presented that the need is strong for the standardization of cloze procedure in terms of the number of errors per grade level so that a cloze score will establish a grade level. Cloze procedure scores are related to reading comprehension scores under a variety of conditions. Cloze procedure has been found to be strongly related to the verbal aspects of intelligence. This would be expected from other sources of evidence. And finally, cloze procedure has been used as a successful reading comprehension teaching device by several investigators. However, it is difficult to establish the differential effects which may be attributable to motivation for other than the cloze procedure materials themselves.

On the basis of this literature the present study presents data on a number of passages of relatively equivalent established readability for pupils of the fifth, seventh, ninth and eleventh grades. In addition the study presents data relating to deletions of varying syntaxes and the reading comprehension of both standardized reading tests and reading comprehension tests built for specific cloze selections.

Third, the study examines an area which is only touched tangentially in the literature, that is to say, the relationship of age and sex to cloze scores.

Fourth, the study examines the relationship of intelligence test scores to cloze procedure scores, as well as reading comprehension scores on the same material.

Fifth, the study examines the effect of presenting a prolonged sequence of cloze procedure materials under control conditions where motivational arrangements of material as well as overt feedback are not presented to the pupils, and where the difficulty of responding is controlled by controlling the syntax of the deletions.

CHAPTER III

PROCEDURE

GENERAL PLAN

The questions to be asked fell into two general categories--those which required a cross-section of several grade levels for relatively short exposure to cloze procedure materials, and a second series of questions which required longer time exposure to cloze procedure materials in a more longitudinal fashion.

The Cross-Sectional Experiment. The cross-sectional experiment involved pupils from the fifth, seventh, ninth and eleventh grades. The number of pupils involved by grade level for each sex is given in Table VI. The pupils in the cross-sectional experiment were involved with the experimental material for three weeks. Six cloze procedure selections were administered to the pupils, two each week on non-consecutive days. The data for the cross-sectional experiment involved the number of correct answers to comprehension questions asked by the various forms of the cloze procedure materials.

The Longitudinal Experiment. The longitudinal experiment was conducted with ninth grade pupils. The pupils were pre-tested with the differential aptitude battery and the California Test of Mental Maturity. For twelve weeks they received two cloze procedure selections delivered on non-consecutive days for a total of twenty-four. This was followed by cloze procedure testing, the language and reading sections of the Iowa Test of Basic Skills, and an information retention test.

Table VI

CROSS SECTIONAL EXPERIMENT SHOWING NUMBER OF PUPILS BY GRADE LEVEL AND SEX

Grade Level	Sex	Statistic	0 Control	1 Random	2 Noun	3 Verb	4 Modifier	5 Prep & Conj	6 Noun Det.
5th	Boys	N	136	81	94	96	75	102	121
	Girls		64	103	136	94	33	82	59
7th	Boys	N	99	97	65	42	67	99	96
	Girls		90	59	59	67	114	73	64
9th	Boys	N	93	98	84	94	64	116	94
	Girls		136	44	43	80	119	61	83
11th	Boys	N	111	78	59	66	76	16	43
	Girls		96	93	92	74	65	96	76

Here student absence became an important variable in the experiment and any student who did not complete eighteen of the twenty-four cloze procedure tests or who was absent for any of the pre-tests or post-tests was eliminated from the experiment.

Population and Sample. The pupils for the two experiments were drawn from a suburban central school in upstate New York, adjacent to a city of 350,000 inhabitants. The population was essentially middle class with a small proportion of rural pupils from nearby farms. The experiment involved a total of forty-nine classrooms in the cross-sectional and longitudinal experiments combined. The experiment involved the total population of classrooms in the school system at the fifth, seventh, ninth and eleventh grades.

TEST MATERIALS

Pre-Tests. The pre-tests administered to the ninth grade pupils involved in the longitudinal experiment were the Differential Aptitude Test and the California Test of Mental Maturity. The Differential Aptitude Test was chosen because it gives a reading score in addition to numerical and spatial relation scores which were to be correlated with the cloze procedure data.

The Post-Test. The post-test consisted of one standardized test and two tests developed specifically for the experiment. The standardized test was the Iowa Test of Basic Skills (the language and reading sections), the language section giving correlational data to relate to the cloze procedure scores in the various syntactical deletions. The reading test scores showed differences in reading ability between groups.

The Final Comprehension Test. The final comprehension test consisted of 144 multiple choice questions in groups of six, which were taken directly from the comprehension questions presented following each of the cloze selections. Since the amount of time since a pupil had read the selection varied in systematic degrees these questions afforded a measure of recall over a period of twelve weeks. This test was administered the thirteenth week. The questions were selected at random from the twelve comprehension questions following each article. Control was obtained for differences between materials and differences in questions as a function of the mixed orders in which children took each cloze procedure selection.

The Cloze Procedure Final. The cloze procedure final was a cloze selection, one thousand words in length, which was designed to determine whether or not there were differences in the ability of individuals to fill in cloze procedure blanks as a function of the training which had occurred from the twenty-four preceding articles. This cloze procedure selection was essentially a random deletion selection which was produced in the same fashion as the cloze procedure material used in the experiment itself. The exception to this was that the number of deletions was balanced according to syntactic type. There were twenty deletions of each experimental syntactic type in the test. This test was graded according to the number of correct responses per type of deletion, and hence differences between the ability of pupils with different deletion type experiences could be determined.

Limitations of the Testing Program. The pre-test administered to the subjects were those tests customarily delivered by the school system. It was felt that additional pretesting would add an undue burden on teachers and pupils. The post-tests, however, were dictated by the needs of the experiment.

Selection. Twenty-four selections of six to seven hundred words in length were used in the experiment. Each selection contained a complete episode in that they were judged to be: 1) understandable without previous knowledge of the area or characters or the subject; and 2) conflicts and issues raised in the material were resolved within the six to seven hundred words.

Readability. The readability level of each of the articles or selections used for the experiment fell within the range of fifth to sixth grade on the Yoakum Readability Formula. On several occasions this required an adjustment of the readability of the article. Typically the readability was adjusted downward. The adjustment was made by substituting words under the following conditions:

1) No proper names were changed;
2) The substituted word must be a synonym of the deleted word;
3) The substituted word was both shorter and more familiar than the original word;
4) If a word occurred in several places and this word was changed, the same substitution was made throughout the selection;
5) There was no change in the denotative meaning of the passage as the result of the change in the words;
6) The connotative meaning was as close to the original connotative meaning as possible;
7) No more than six different words were changed in any single selection.

These procedures allowed the construction of twenty-four separate passages of equivalent readability without a change in the meaning or distruction of the author's intent. Such a procedure was essential since the experiment was designed to compare each of the twenty-four exercises against the other exercises as they were presented sequentially to the children. Since the sequences of presentation to the students were essentially random, a deviation of readability would make the scores not comparable.

Comprehension Questions. Since it was the intent of the experiment not only to examine the cloze procedure per se but also to examine its effect upon the reading comprehension of material, twelve comprehension questions were developed for each of the selections. Six of these questions were factual and six were inferential. The questions were multiple choice with four options. The correct response position was assigned at random. The factual questions were questions about material contained directly in the selection. The inferential questions required that the respondent infer from the material in the selection a most probable answer.

Pre-Testing the Questions. The quality of the questions and the ability of the questions to discriminate was important for the study. Each selection and its comprehension questions were pretested with seventh-grade students under the following conditions:

> The same class of students was used for pretesting all the materials. The pupils were a single class of seventh grade pupils in the junior high school of a small Connecticut city with a population of about 20,000. The class selected was of average ability, being the middle seventh grade section of an

ability-grouped class. The materials were presented to the students during a reading class. Two selections were presented each week on non-consecutive days for twelve weeks. The pupils were instructed to read the selection as carefully as possible and to answer the questions which followed them. The data was collected on mark-sense cards and processed through the test scoring program on a 1620 computer. The 1620 computer provided scores for each of the students, and an item analysis which was used to alter the comprehension questions and to control for difficulty with the following restrictions:

1) Each option or distractor must draw at least one response;
2) The number of correct responses must be between sixty and seventy percent, to account for the fact that some of the materials are intended for fifth, ninth and eleventh grade pupils as well as seventh grade pupils;
3) The items must discriminate between better and poorer students significantly

All items not meeting the above criteria were rewritten and the selections retested with another group of seventh-grade pupils with similar characteristics to the first, from a suburban township in central Connecticut.

Deletions. Once the materials had been pre-tested and the questions met the above criteria, each of the twenty-four selections was prepared in six cloze forms and one reading form. The cloze forms were prepared by deleting every tenth word, starting with a randomly-assigned first word. It is typical to use a twenty percent deletion in cloze materials; however, a ten percent deletion was used for the following reasons:

1) Experiments by Taylor and others show that deletions of ten or twenty percent make relatively little difference in the scores of students;

2) If the twenty percent level was chosen, it would be virtually impossible to get equal numbers of deletions in the various grammatical categories. The deletions were made according to the following procedure for each of the forms:

Form 1. Reading Selection: No words deleted.

Form 2. Random Deletions: Every tenth word, starting with a randomly selected first word, was deleted regardless of syntax.

Form 3. Noun Deletions: Ten percent of the passage was deleted. Only nouns were deleted. Proper names and pronouns were not included in the deletions. The initial word was chosen by random selection. If the random selection was not a noun the nearest noun to it was chosen. The noun nearest the tenth word following was likewise deleted. If two nouns were equidistant from the tenth word which was not a noun, the noun which occurred last was deleted.

Form 4. Verb Deletions: Ten percent of the passage was deleted as with the nouns. The first word was chosen by random selection. If this was not a verb, the verb nearest to the word was deleted. The rule for succeeding deletions was as above. Gerunds and auxiliary verbs were excluded from the deletions.

Form 5. Modifiers: Ten percent of the passage was deleted in the manner described above. The ten percent included adjectives and adverbs.

Form 6. Function Words--Prepositions and Conjunctions: Ten percent of the passage was deleted in the manner described above. The deletions were all prepositions and conjunctions.

Form 7. Function Words--Noun Determiners: Ten percent of the passage was deleted in the manner described above. The deletions were all noun determiners.

Other Grammatical Categories: Two other grammatical categories were attempted--auxiliary and functional verbs, and pronoun substantives. In neither case did they occur in sufficient number to allow ten percent deletion in all twenty-four selections.

GRAMMAR

A number of problems present themselves to any investigator who applies the cloze technique to some sort of grammatical schema. As it has heretofore been used, the cloze technique involves deletions of single words and depends for its effectiveness as a measure on the degree of difficulty of replacing a single word, a variation which may be regarded as posing the problem of filling a smaller or larger closure in prose. To delete more than one word without designating the gap as a complex problem in which more than one word is missing makes the task enormously more difficult for the reader who can in many cases fill a given blank with a whole nominative or verbal construction. Numbering two or more consecutive blanks to indicate gaps of more than one word reduces the difficulty somewhat but the reader is still likely to be misled. One of the chief problems, then, is to build a design which takes as its basic unit of measure the word in a language not always measurable in such units. For example, New York and put up with stand for a particular place and a particular act, and to ask which word is the noun or which word the verb in these multiple-word units is to ignore the way they function in the language. Among function words, one observes that three words, in spite of, can be replaced with despite with negligible loss of meaning, but deletion of despite is likely to frustrate the reader unnecassarily even when he knows what connective should be substituted--the three words he uses more commonly to express this connection, in spite of, will not fit a one-word blank. The investigators found, particularly among readers in the lower two grade levels examined, that there was an occasional tendency to fill in blanks with multiple word constructions that made grammatical sense and in some cases showed a considerable sophistication in reading and digesting the general information conveyed in the context of the prose.

By the same token, contractions such as that's and there's and I've, which may be designated by a single blank, do in fact require a complex judgement by the reader, and, like despite, provoke a frustration in the reader who more regularly sees an actor and an action represented in two words. A third problem is posed by separable verbs, which may occur far enough apart from each other to avoid the difficulty of consecutive word deletion, but once again are a single expression denoted as two word units.

The traditional grammatical scheme of parsing sentences has the faults of any system which uses the interrelationships of one language to explain those of another very different tongue. Though some of the classifications of the older system are sound many more, particularly the parts of speech, are both unwieldy and misleading. Pronouns, for instance, are of various kinds and functions, and a description such as "a word that takes the place of a word designating a particular person, place or thing", must be stretched considerably (even when one understands all the grammatical and lexical meanings in an utterance) if we are to account for demonstrative, genitive, interrogative and other uses of words so designated.

It was not, however, the purpose of the investigators to account for all of the language, but rather to isolate every class of words which constituted at least ten percent of the words in running prose so as to examine their individual contribution to the meaning of reading passages and to test the effects of their loss on grammatical and total meaning. This meant that some kinds of words were omitted from consideration, and because of difficulties already discussed some examples within classes were regarded as unusable in the design of the experiment. Among nouns and verbs the traditional definitions were used though the function of individual examples was arrived at in a slightly different manner than the traditional methods of parsing. There was then a regular attempt to maintain consistency within a given class of words. Black dog and radar operator are not parallel constructions, though the first word in each pair is commonly called an adjective, and even such a recent study as that of Ruddell[32] makes the older description the basis for regarding radar as a modifier in such constructions. The methods of identification of classes is based upon W. Nelson Francis' The Structure of American English and Charles Carpenter Fries' Structure of American Grammar. These then were the classes of words deleted.

Nouns. The traditional definition of a word that indicates some sort of object, place name, or person was the guideline. In cases where the mechanical system of mutilation pointed to a word which was itself a noun but which was one functioning unit of a multiple word nominative, e.g., Brooklyn Bridge, the word was skipped and the closest single-unit nominative was deleted instead. Nouns were identified by a number of grammatical cues. Noun markers such as a, the, some, any, few, his, regularly signal a following noun even when other words intervene between the marker and the noun. Where morphology indicated that a given word might fall in one of two or more classes, syntax was taken as the grammatical clue. Thus, "the last man," "a friendship which lasts," and "this home was his last" are differentiated by one or more of these methods. Both proper and common nouns were deleted.

Verbs. This class was restricted to specific verbs, that is, function verbs such as do and have in constructions such as "do drop us a card," and "have you met Miss Jones?" were excluded.

An attempt to construct cloze tests from which function verbs and auxiliary verbs were deleted proved impossible for a large number of extended reading passages of various types. The number of words of this class fell far short of the percent necessary for the deletions, and attempts to write or select passages with the necessary number of such verbs were rejected because readability was high, the interest level or appeal of the passages for a variety of students was low, and the passages were predominately written in the historical past, precluding the variety the investigators felt necessary for a general appraisal of the reading process.

Specific modifiers. This group included both adjectives and adverbs, precluded the relatively small and fixed body of words represented in the above examples of noun markers, and verb markers. This group was differentiated by means of the inflected endings, e.g., -ly, -er, -est, and the syntactic relationship to words of classes one and two.

Prepositions and conjunctions. This body of words included all prepositions and conjunctions except so-called "subordinating" conjunctions, words which are by themselves relatively empty of lexical meaning but do in fact point to relationships among other lexical units.

Noun determiners. This class comprised the articles a and the and also included certain possessive, demonstrative and interrogative pronoun uses (some of which are noted above as "noun markers") in such pronoun uses with nouns, e.g., "that man," "whose glove?", "what gall!" and "his mind," as well as some, any and few when they exercised a determining effect over the nouns following them. The group is, in fact, another body of "structure" words which do point out structural meanings and interrelationships among sentence constituents but exist because of their grammatical meaning rather than a lexical meaning.

Pronoun Substantives. The last class of words attempted included the substantive uses of pronouns when syntax indicated that they did not fall into the class designated "noun determiners" and where the individual words stood as complete markers for a nominal unit already evident, whether that nominal unit was one word or a noun clause or the whole thought denoted in one or more of the previous paragraphs. It also included expletive uses such as there in "there was a crooked man..." The following four examples of contrastive utterances will serve to illustrate the general method by which words of this class were differentiated, even though the same words were also categorized as noun determiners and they all occur in what are generally called the "subjects" of the sentences.

Noun determiners: Some day my prince will come.
Many great hearts are asleep in the deep.
Few men are called.
Such men are true patriots.

Substantives: Some is better than none.
Many are called.
Few are chosen.
Such, were the joys.

Pronoun substantives did not occur in the sample to a sufficient degree to warrant inclusion.

Perhaps for the more sophisticated student of reading and writing, the most important category of words omitted from these classes is the body of logical connectives represented in such words as consequently, moreover, however and therefore. A test of these function words which link a logical succession of utterances would probably increase our understanding of whatever mental shorthand the reader uses in comprehending prose if we could construct a method for deleting a sufficient number of them to contrast with words of other classes. But again the investigators were confronted by an important class of words which constitute a very small percentage of the total number of words in running prose and tend to appear most frequently in rather difficult non-fiction.

While this scheme of classification is by no means an attempt to create an adequate grammar of American English it does describe the functions of more than ninety percent of the words in the running prose examined. It will not begin to describe much of the structure of many modern colloquial works or non-standard dialects. It does provide a convenient and systematic method for deleting words of various kinds.

ORGANIZATION OF MATERIAL FOR PRESENTATION

To control for order of presentation effect and possible differences in the materials, the twenty-four selections were randomly assigned to four blocks of six selections each, labeled A, B, C and D. Selections were presented to the students in blocks of six. For the longitudinal experiment the first presentation thus included all four blocks of material. These blocks were reassigned to different students on the second, third and fourth presentation of the longitudinal study.

The longitudinal study was divided into four presentations of three weeks in length. The cross-sectional experiment was three weeks in length. The selections within each block were randomly assigned in sequence of presentation as a further control for order effects. The assignments within block sequence of selection was identical for all forms of deletion. The tables below show the analysis of variance between the blocks of material indicating no significant difference between the blocks and thus allowing grouping of the data between blocks for final computations.

Table VII

MEAN CORRECT COMPREHENSION QUESTIONS BY BLOCK AND GRADE

Grade Level	Block A	B	C	D
5	5.64	5.20	5.04	6.73
7	5.85	6.46	6.39	7.90
9	6.98	8.40	6.60	8.78
11	7.96	8.64	8.75	6.98

Table VIII

ANALYSIS OF VARIANCE OF BLOCKS BY GRADES

	Sum of Squares	d.f.	Mean Square	F
Rows (Grade Level)	14.33	3	4.78	6.09
Column Blocks	2.26	3	.75	.94
Residual	7.17	9	.80	
Total	23.76	15	1.58	

$F_{.95}(3, 9 \text{ d.f.}) = 3.86$

$F_{.99}(3, 9 \text{ d.f.}) = 6.99$

The cross-sectional experiment was three weeks in length, and each student was assigned a single block of material. The students were assigned to blocks of material at random.

Assignment of Pupils. Each pupil in the experimental classes was assigned a block order and a deletion form by random procedures. Thus the experiment was replicated within each classroom. Pupils were randomly assigned to treatments and orders. This procedure had the additional benefit of making it impossible for one student to use the responses of another. Each block of material was inserted in an envelope labeled with both its block and its form number. The envelope contained the following materials:

1. Six selections, each with the same deletion form, and coded with a presentation number;
2. Six sets of twelve questions, one set attached to each of the appropriate selections;
3. Six IBM mark-sense cards coded to match the selections with the presentation order number. The mark-sense cards were pre-punched with the selection order number, deletion type, and student number;
4. One IBM mark-sense name card to provide student name, age, grade and sex, coded as above;
5. Six answer sheets with spaces for 75 answers, as well as the name, age, sex, grade, student number, and selection number.

To limit outside interference and the possible Hawthorne effect of having observers in the room, as well as to provide for generalization to other classroom situations and to randomize the experimenter variable, the study was supervised by 59 classroom teachers. The envelopes containing the above materials, with the students' names written on them, were delivered to the teachers.

The teachers were instructed to:

1) present the material to the assigned pupil at the beginning of class, two periods a week, on non-consecutive days;

2) specify the material for the day by number, and be sure that the pupils used the correct IBM card;

3) be sure that the students put their names and other data on the answer sheet

4) allow as much time as necessary to complete the material;

5) insure that the students answering the comprehension questions did not refer back to the selection or to their answer sheets;

6) provide seat work for the pupils when they finished their material;

7) have the materials replaced in their envelopes and returned to the teacher at the end of the hour.

Pupils' Instructions. The pupils' instructions were read by the teacher and were as follows:

"The following exercises will help you in your understanding of the English language. From your envelope remove Test No. ___ and the packet of answer cards.

"Now, find the brown name card and the answer card with the same number as the test on it. Fill in the name card with your name, age, sex, and grade.

"Now, remove one answer sheet from the envelope, fill in the top lines with your name, age, sex, grade, test number, and the date.

"Be sure your answer sheet and the IBM card have the same number as the exercises. Are there any questions thus far?

"All of the exercises you students have are different. Some of them will have no blanks. In this case read the exercise carefully and proceed to answer the comprehension questions.

"You are to read the exercise. Wherever you find a blank, write the word which should go in the blank on your answer sheet. Be sure to write the word opposite the number on the answer sheet which corresponds to the number of the blank on the exercise. Work as carefully as you can. You will have enough time to

> finish. When you have finished the exercise, turn to the comprehension questions which follow. Answer these by making an appropriate mark on your IBM card with the special pencil. Do not look back at the exercise but answer as well as you can from your memory. When you finish, raise your hand. I will collect your materials. Are there any questions?"

SUMMARY

Thus the treatments and orders of selection were randomly assigned to students, who filled in the missing words in the exercises on separate answer sheets and answered the comprehension questions during a regular class period under the supervision of a classroom teacher.

TREATMENT OF THE DATA

A total of 1411 students participated in the two experiments. In the longitudinal experiment a number of pupils did not complete all twenty-four exercises due to lack of ability, boredom, or absence. The students who did not complete at least eighteen of the exercises were dropped from the data analysis with the reasoning that any influence from sustained cloze procedure could not be detected with fewer exposures. There were 9,198 completed sets of data on both the longitudinal and cross-sectional experiments. These included both the answer sheet and the mark-sense IBM card. Two whole classes were dropped from the experiment because the instructor felt the material too difficult. Both of these classes were low-ability groups at seventh grade level. The total N for the longitudinal experiment was thus 114. The _N_ for the cross-sectional experiment was 976.

The data cards for each student were cleaned of stray marks and the marks made by the students were darkened prior to processing in the IBM 519 to punch the mark-sense data into the data cards. These cards were sorted by student namber and by deletion type as well as by order of presentation. The cards were scored on the 1620 with a test scoring program provided by the Cooperative Test Scoring Service of the University of Connecticut. The data thus obtained gave the number of correct responses for each of the twenty-four tests for each student, as well as a discrimination index for each item. The cloze answer sheets were hand scored by student assistants to determine the number of correct responses.

These data were then transferred to tabulation sheets along with the test scores for both pre- and post-tests. The data were punched on cards and processed on the 7040 with a program which produced means, variances, standard errors, and correlations of all the variables. A second program was used to produce t-tests between each variable for each form. The correlation matrix thus produced was 111 x 111. One of these was produced for each of the deletion types.

Two pilot studies were conducted. These are reported on the immediately following pages in Chapters IV and V.

CHAPTER IV

PILOT STUDY I

The cloze procedure has been shown to produce gains in reading comprehension. When three-hundred-word prose selections were given to college students, results on comprehension tests showed that readers of cloze passages with ten percent of the words deleted displayed significantly greater grasp of the material than readers of the same passages with all the words (Bloomer, 1962). Previous studies have also shown that when words are deleted at random, say every nth word, there is a difference in the difficulty of replacing words of different grammatical classes, e.g., modifiers, nouns, verbs, prepositions and conjunctions, and articles (Rankin, 1958). The present study, then, is concerned with the effect upon the immediate comprehension when cloze deletions are made not at random but uniformly by grammatical class.

Since the immediate reading comprehension score seems to be related to the final gain made after the reader has done a number of cloze procedure exercises, it is important to determine if some classes of words tend to destroy the structure of prose and cause loss of meaning, and which elements increase the probability of gains in comprehension.

Procedure. On June 3rd and 4th, 1964, 1148 fifth and seventh grade students of the schools of South Windsor, Connecticut were given a single reading passage in seven forms followed by a single comprehension test of twelve questions. The prose was the initial section of a a short story by author-dramatist William Saroyan. To retain the artistic aims of the writer in the piece of fiction as a whole, the prose was edited only slightly to produce a shorter self-contained unit of fiction and to reduce the original readability from 6.00 (on a modification of the Yoakum scale which omitted common contractions from consideration) to 5.00.

The short story is a narrative made up largely of dialogue between a twelve-year-old boy and an Indian friend who does not speak the elided language of the popular prototype Indians of television and movies.

Six of the forms of the passage were modified so as to produce an evenly distributed deletion of ten percent of the words. As a control, a seventh form was left unmutilated. The six kinds of deletions were 1) random; 2) nouns, proper and common; 3) verbs, exclusive of auxiliary or function verbs; 4) specific modifiers--adjectives and adverbs; 5) prepositions and conjunctions, exclusive of subordinating conjunctions; 6) noun determiners--e.g., articles, possessive pronouns, words of the type represented in some, any, few, which precede nouns and exercise a modifying effect over them. All deletions in a given form (except the random form) were of a single word class.

The students were instructed to read the passage filling in any missing words on an answer sheet provided, and then to answer the questions following the selection on a mark-sense card to be scored by machine. The first six of the questions tested comprehension of the literal meaning of the passage. The last six were designed to elicit inferences logically implicit in the passage. No time limit was imposed.

Results. The scores on comprehension tests (see Table IX) show that readers of passages with noun determiners deleted had a greater understanding of the passage than all other groups, including those who read unmutilated prose. There was a significant loss of comprehension among those who read passages with specific modifier deletions. The scores on comprehension tests were (from high to low):

1) noun determiners;
2) no deletions;
3) prepositions and conjunctions;
4) modifiers;
5) specific verbs;
6) random;
7) nouns.

Noun determiners were superior to all other classes of deletions and to the control group. The no-deletion control group was superior to all deletion types excepting noun determiners. In turn, prepositions and conjunctions were superior to modifier, specific verb, noun, and random deletions.

In the number of exact replacements of deleted words (see Table X) scores reveal that prepositions and conjunctions, and noun determiners, were more regularly replaced than any other class of word. Scores on word-replacement tests ranked from high to low as follows:

1) prepositions and conjunctions;
2) noun determiners;
3) random;
4) verbs;
5) nouns;
6) specific modifiers.

Conclusions. Marked losses in comprehension occur where the structure of the prose and its meaning are broken by deletion of the basic meaning carriers of the language, i.e., nouns, verbs and modifiers.

Table IX

MEANS AND 't' RATIOS OF COMPREHENSION QUESTIONS CORRECT

	0 Control	1 Random	2 Nouns	3 Verbs	4 Modifiers	5 Prepositions & Conjunctions	6 Noun Determiners
$\overline{X}$	6.89	6.88	6.69	7.02	5.14	7.38	7.64
N	58	50	55	49	50	52	50
Control		NS	NS	NS	3.95	NS	NS
Random			NS	NS	3.91	NS	NS
Nouns				NS	3.78	NS	2.41
Verbs					4.27	NS	NS
Modifiers						5.89	6.10
Prepositions & Conjunctions							NS
Noun Determiners							

Table X

MEANS AND t's OF PERCENT CLOZE BLANKS CORRECT (All Students

	Random	Noun	Verb	Modifier	Prep.& Conj.	Noun Det.
$\bar{X}$	63.57	42.75	55.41	39.95	73.80	70.86
N	50	48	49	49	48	42
Random	-	6.21	2.63	7.29	3.26	2.67
Nouns	-	-	3.98	NS	9.70	10.00
Verbs	-	-	-	5.04	6.21	6.13
Modifiers	-	-	-	-	10.95	11.58
Prepositions & Conjunctions	-	-	-	-	-	NS
Noun Determiners						

It is apparent that even in passages of fairly low readability level the duplication of lexical and structural meaning is not great enough to supply the contextual information necessary to replace these words and recreate the sentences as the same meaningful utterances they were before mutilation. Among the words deleted at random, forty percent of the words deleted fall into these three classes, but only nineteen different words are presented, eleven of which appear at least one other time in the passage. Both the number of words of these three classes compared with the total number of words and the number of words that appear only once (and therefore never appear when deleted) would seem to account for the gain over the control group. Of the remaining sixty percent of the deletions in this form, all but two were function words which contribute no lexical meaning of their own but relate parts of utterances. These words, among the commonest in English, are dictated by the grammatical relationships of the sentence signals which are regularly duplicated by the morphology and syntactic patterns of the language. Significantly, one-fourth of the structure words deleted were noun determiners, which may account for the perceptible gain among readers with random deletions. Among the selections with verbs deleted, the missing words are all among the commonest four thousand words in the language, according to the Thorndike table, and forty of the sixty are among the commonest two thousand words. There are, in fact, only thirty-three different verbs represented in the deletion, which points to a lexical duplication of significant proportion. This would account for the lack of perceptible loss among readers confronted by passages with missing verbs.

The gain among readers with noun determiners deleted can be explained in a number of ways, and several factors may well be operating simultaneously to produce the increase.

1. Focus. Removal of function words where syntax and morphology already signal the grammatical relationships in a sentence may tend to focus the attention of the reader on the larger meaningful units of the passage being read. This would be particularly true of noun determiners, the sequences of which are governed not only by the sentence in which they occur but by the sentences (and paragraphs) which precede them. Moreover, if we grant Benjamin Lee Whorf's assertion that the bulk of the abstract words in English are nominative, the increased reader attention to words following marked deletions would seem to be a focusing of the perceptive powers on the individual elements where gains and losses in meaning occur. This would be borne out by the lower scores following noun deletions. Since the other body of function words (prepositions and conjunctions) would also tend to produce structure-centered reading as well and yet not concentrate attention on nominative elements, the difference in comprehension scores between classes of function words would tend to confirm the hypothesis that nominatives carry the burden of abstract meaning. Where the larger meaningful structures are shattered by deletion of words which cannot be supplied, e.g., specific modifiers, it is likely that the reader shifts his attention to the less meaningful units. Since deletion of prepositions and conjunctions also produced higher comprehension scores than all other forms of deletions except noun determiners it is plain that a structure-centered reader is more likely to read with comprehension.

2. Novelty. The requirement that the reader substitute as accurately as possible the missing word provokes in him a response to a novel reading experience, and when the reader is reasonably confident that he has correctly replaced the missing word--highly probable where he is supplying function words dictated by the grammar of his language--this may tend to relieve other anxieties he may have when subjected to the reading-testing ordeal. Correspondingly, if he is unable to assure himself that he has replaced the words correctly (as in cases where he must guess at specific modifiers) his comprehension of the material may be affected adversely by a frustration factor.

3. Repetition. Repeated reading required to supply missing words, particularly where the words are governed by structural principles not always apparent in the immediate sentence, would appear to be a form of reinforcement of the material. Repeated reading where the missing link is incomprehensible would tend to break down the unity of the material under scrutiny, reinforcing the baffling aspects of the material and solidifying in the reader's mind only that aspects of the material are beyond his grasp.

4. Time. The extra time required to supply missing words in the text would increase the exposure of the reader to the material, allowing more time for the information to crystallize, and the necessity of stopping regularly to supply the missing words would constitute a regular break, not offered by paragraphing and punctuation, thus providing a method of deliberately laminating the information in layers which are fairly uniform. The extra time where the words are not readily supplied would allow the reader to reinforce only his knowledge that there are significant gaps in his understanding of the material. He may, in fact, spend so much time attempting

to place the word that he loses what fragments of information are presented to him.

It is apparent that if structural relationships within an utterance are expressed simultaneously by two or more grammatical devices, one or more of these structural pointers may be unnecessary. Morphology and syntax by themselves, however, may be insufficient to convey the function of a word and thus the relationships among several words. When too many of the function words are missing the result is likely to be ambiguous if not meaningless. Professor Fries' now classic example "Ship sails today" indicates one result of too few grammatical signals.

Since lexical redundancy varies markedly from one prose sample to another, we may not infer that deletion of verbs and nouns produces no loss in comprehension in all kinds of prose. And it may well be that the number of different words within a given class when several classes are deleted at random will affect the reader's comprehension of the passage.

The present study in addition to particularizing the general findings of the earlier experiment with three-hundred-word passages allowed the experimenters to test the effects of the longer passages, using a fairly large number of words without increasing the percentage of deletions. The longer passages are perhaps more realistic appraisals of general demands made of readers. A future experiment will disperse the deletions to an even greater extent among various kinds and modes of literature, involving subliteral, e.g. symbolic or ironic, meanings.

CHAPTER V

PILOT STUDY II

It is probable that when materials are used to elicit cloze responses, all of which are function words, the subscript line marking a missing word is a variable component of the reading processes. Function words, or "functors," are by definition members of a body of fixed units in language which must be memorized because their meanings are chiefly grammatical. They contribute relatively little lexical meaning to an utterance, and stand as markers of the interrelationships of other constituents in an ordered stream of words. Because this small, fixed body of words is committed in constructions by one who learns English and because such words compose about one third of the bulk of the language as it is spoken, the presence of any marker in many constructions is likely to provide the probable link in grammar.

It would also seem that the requirement that the reader make an actual replacement of the word by writing it in should affect the reading process by interfering with the continuum of a prose passage, perhaps producing considerable rereading of individual segments. In the case of function words which earlier studies indicate are supplied readily, the investigators speculated that perhaps the subscript marker supplied sufficient syntactic information for the reader to grasp grammatical relationships and that the regular stops to write in words tended to fragment the prose for the reader.

To assess the effects of both the subscript marker or underscore and to appraise the effects of requiring a reader to write in his cloze responses, a reading selection was prepared in several forms and assigned at random to 420 seventh and eighth grade students of two Manchester, Connecticut schools.

Procedure. Words deleted were of three classes, one class to a subject. An evenly dispersed deletion of ten percent of the words (all of a given class) was obtained by dividing the prose passage into ten-word stiches and choosing the word of the desired class as near the end of each stich as possible.

Class one was prepositions and conjunctions, exclusive of subordinating conjunctions.

Class two was noun determiners, also known as "noun markers", such as a, the, some, any and possessive pronouns in constructions such as his book.

Class three was a body of pronoun substantives which, occurring alone, replaced a larger or more specific nominal unit already in evidence.

A single passage was mutilated to prepare cloze materials of three types, so that the passage appeared in nine mutilated forms and an unmutilated control form.

Type A, comprising three sets of deletions (Classes one, two and three) was a standard cloze form followed by a twelve-question comprehension test. Where words were deleted a five-pica subscript line appeared with a number corresponding to numbered blanks on an attached answer sheet. Subjects were asked to fill in the missing words by writing them on the answer sheet, and then to answer the questions without referring to the passage.

Type B (Classes one, two and three) was the same as A except that the answer sheet was eliminated. Subjects were asked to guess the missing words indicated by the subscript lines as they read but were not required to write the words and they were asked not to write on the test booklet. When they had finished they were to respond to the comprehension questions without consulting the reading passage.

Type C (Classes one, two and three) passages were not standard cloze passages. Ten percent of the running word count was absent with no symbol to indicate that such mutilation had been made. In Type C, form one, for instance, the same prepositions were absent from the same contexts as in form one of Types A and B, with the difference that no subscript line indicated the omissions. Subjects who received Type C forms were told nothing of the omission, but instructed only to read the passages and to respond to the comprehension questions following their reading without consulting the passage again.

The passage was the same one that had been used earlier on fifth and seventh grade students at South Windsor, Connecticut, schools, a six-hundred-word narrative of a twelve-year-old boy and an Indian friend. Readability measured by the Yoakum scale was 5.0. The passage was colloquial in nature, and the unmarked deletion of function words in many places produced pidgin dialect not unlike the ordinary speech of motion picture Indians.

The passage was followed by twelve multiple choice comprehension questions which had been pretested in the manner described in the procedure. Four hundred and twenty seventh and eighth grade pupils were assembled in two sessions in a junior high school for the purpose of the experiment. The subjects included all of the children in seventh and eighth grade who were present in school on the day of the experiment. The subjects were randomly assigned to treatment group and to class of deletion. All treatments were given simultaneously to all children, thus avoiding any effects of selection of children for school class, sex, or intelligence. The pupils were given appropriate written instructions to follow.

Results. The data for the present experiment were the mean number of comprehension questions answered correctly by pupils in the various treatment groups. These data are found in Table XI. Data were treated by analysis of variance as shown in Table XII, and indicate that no significant difference was found between either methods or classes of deletion.

The results of the experiment therefore indicate that the subscript line or the requirement to fill in the missing word in cloze procedure does not produce a result which is different from reading material from which words have been deleted but where there is no indication other than syntactic or structural indications of the deletions. Whatever effects therefore may be found in the cloze procedure will appear to be not a function of the physical structure or the requirement to make responses, and any short-term results with non-reinforced cloze procedure therefore must be seen as a function of the language mutilation in and of itself and the information which is carried through the mutilated language.

Comparison of this group with a similar population in the main group indicates that each of the three treatment types is superior to the control group for preposition and conjunction deletions, and for noun determiner deletions. Any of these forms of handling a single cloze exercise seems to produce superior comprehension than does straight reading or mutilated reading at the seventh grade level, though a specific test of this hypothesis would seem warranted.

Table XI

MEAN NUMBER OF CORRECT COMPREHENSION QUESTIONS BY TREATMENT AND DELETION TYPE

	Treatment Type		
Deletion Type	1 Cloze Non-Answers	2 Cloze Answers	3 Pidgin English
Prepositions and Conjunctions	7.6	7.9	8.3
Noun Determiners	8.2	8.3	8.1
Pronoun Substantives	7.8	8.5	7.8

Table XII

ANALYSIS OF VARIANCE OF MEAN COMPREHENSION QUESTIONS CORRECT BY TREATMENT AND DELETION

	Type Sum of Squares	d.f.	Mean Square	F
Rows (Treatments)	.20	2	.1	1.03 N.S.
Columns (Deletion Types)	.11	2	.055	.56 N.S.
Residual	.39	4	.975	
Total	.70	8		

CHAPTER VI

THE EFFECTS OF NON-OVERT REINFORCED CLOZE PROCEDURE UPON COMPREHENSION BY GRADE AND SEX

If one were to follow strictly the tenets of operational behaviorism, the process of reading is an extinction process. No overt reinforcements are given for reading words as sentences except at early grades, unless we consider that such overt behavior as page turning or finishing a book is in fact sufficiently reinforcing to sustain the process, or if some extrinsic award is awaiting the completion of the reading. While it is true that some derive considerable reinforcement from having finished a book and then reporting to friends or teachers that it has been read, the vast bulk of reading is not for this purpose. While this is an oversimplification of the reinforcement position, it points to the necessity for the postulation of a non-overt reinforcement system to account for the phenomenon of sustained reading. Self reinforcement is derived from the individual's ability to match the author's language patterns with his own to derive meaning from the printed word and to interpret these meanings in view of his own experiences.

The ability of an individual to derive self-reinforcement from reading material is dependent upon the magnitude of the reinforcing stimulus, which in turn is directly related to the difficulty of the reading material. Where material is too difficult we find that the ability of the individual to self-reinforce is lower than the amount of energy or frustration that he must exert to read. Alternatively, if the material is too easy, the increment of reinforcement is too small and exceeds the extinction ratio. The ability of an individual to reinforce himself while reading material is directly related to his attention to the material and to his reading comprehension.

Given this position, therefore, we can deduce that reading comprehension will be relatively low where the material is difficult for a reader. In turn, as the relative difficulty decreases, reading comprehension will not increase proportionately. The area of extreme difficulty in reading comprehension has been explored through the development of the readability scales. The area of supreme simplicity, on the other hand, remains to be explored.

The cloze procedure has a tendency to make reading materials more difficult and as such we should expect responses to the cloze procedure to follow a similar pattern but to give different values at different grade levels. Cloze procedure has still a further advantage in that one can relatively control the amount of difficulty of cloze procedure passages by controlling the syntax of the deleted material. Noun, verb and adjective deletions tend to be more difficult, whereas pronouns, conjunctions and noun determiner deletions tend to be considerably simpler. Hence cloze procedure offers a flexible tool for investigating the limits of reading comprehension in relation to reading material difficulty.

The present cross-sectional experiment was designed primarily to compare boys with girls at the fifth, seventh, ninth and eleventh grade levels, and the long-term recall of cloze procedure materials with deletions of various syntax against straight reading material and random cloze.

All cloze materials in the present experiment were ten percent deletions, primarily because it is relatively impossible to use a higher percentage of deletions and at the same time specify the syntax of the deletions.

PROCEDURE

Materials. Twenty-four six-hundred-word passages were prepared. Twelve comprehension questions were prepared for each one. These were pretested and the data analyzed for item discrimination. Adjusted questions or paragraphs were pretested a second time to ensure that the items did discriminate. All selections were adjusted to fifth grade reading level using the Yoakum formula and all the adjustments were made in such a way as not to affect the meaning of the paragraph. Each selection was prepared in seven different forms:

Type 0 - Zero deletions
Type 1 - 10% deletions: Random
Type 2 - 10% deletions: Nouns
Type 3 - 10% deletions: Verbs (auxiliary verbs not included)
Type 4 - 10% deletions: Adjectives
Type 5 - 10% deletions: Prepositions and Conjunctions
Type 6 - 10% deletions: Noun Determiners.

The sample for the present experiment consisted of fifth, seventh, ninth and eleventh grade pupils, totaling 976, from a regional school adjacent to a large city in upstate New York.

Method. Each student was given a package of six of the selections along with the comprehension questions and appropriate answer sheets for answering the cloze material. The envelopes were passed out to the pupil at the beginning of the class period on two non-consecutive days during three weeks. At the end of each working period the materials were collected from the student and remained with the teachers. The student was instructed to read the selections, to fill in the blanks where appropriate, to answer the questions and to return the material to the teacher. The teacher did not grade the material, reducing the possibility of overt reinforcement. Each student received six selections with the identical deletion type but within any class deletion types and selections were randomly assigned to students to control for class differences.

RESULTS

Between-Grade-Level Comparisons. Table XIII gives the means and variances for each deletion type at each grade level. Reading comprehension scores for Grade 7 are superior to those for Grade 5 beyond the .01 level of confidence for all comparisons except for Type 2, noun deletions. The difference is not significant although it is in the expected direction. The t tests range from 2.91 for Type 3, verb deletions, to 6.87 for Type 6, noun determiners.

The comparisons between Grades 7 and 9 indicate that six of the comparisons are in the expected direction and significant beyond the .01 level of confidence. The t's range between 4.07 for random deletions to 6.87 for the zero deletion group. The exception to this trend occurs in the noun determiner group where the reading comprehension following noun determiner deletions for Grade 7 is significantly greater than the Grade 9 comprehension for the same deletion type. Grade 7 is significantly superior at the .05 level of confidence: t = 2.44.

The Grade 9 - Grade 11 comparisons indicate no significant difference between the groups except that Grade 11 is significantly superior to Grade 9 for the noun determiner group, Type 6 (t = 3.31).

Table XIII

MEAN CORRECT RESPONSES TO COMPREHENSION QUESTIONS BY GRADE LEVELS AND TYPE OF RESPONSES FOR ALL SUBJECTS

Grade Level	Statistic	Deletion and Type 0 Control	1 Random	2 Nouns	3 Verbs	4 Modifiers	5 Prep. & Conj.	6 Noun Determ.
5th	N	200	184	230	190	108	184	180
	$\bar{X}$	5.70	4.82	5.16	5.21	4.56	5.68	5.87
	S^2	5.24	5.65	9.49	5.72	5.34	4.88	4.71
7th	N	189	156	124	109	181	172	160
	$\bar{X}$	6.61	6.23	5.28	6.03	5.58	6.88	7.58
	S^2	6.07	5.65	11.38	5.32	6.71	5.49	5.76
9th	N	229	142	127	174	183	177	177
	$\bar{X}$	8.19	7.33	6.83	7.07	7.16	7.98	6.84
	S^2	4.74	8.35	7.67	6.68	4.65	5.55	6.59
11th	N	207	177	151	140	141	112	119
	$\bar{X}$	8.22	7.02	6.84	7.45	6.87	7.66	7.69
	S^2	4.81	5.74	4.64	8.09	2.86	4.58	3.50

It should also be commented at this point that the noun determiner group was not superior to the Grade 7 noun determiner group. Two possible explanations offer themselves for this phenomenon. The first is that the ninth grade reached the upper ceiling of the test and therefore the eleventh grade could not be expected to achieve higher scores. This, however, appears improbable for several reasons. First, the mean number of responses made by eleventh graders was less than two-thirds of the number of possible responses. Further, in four of the comparisons the responses made by the ninth graders was superior to the eleventh graders, indicating that it is possible to get higher scores, albeit these differences were not significant.

It is important to bear in mind that the material which was presented to the eleventh grade pupils, the same as the other material, was at fifth grade level. Hence its level of difficulty for these pupils would be relatively low. As the level of difficulty becomes extremely simple, the value of making a correct response is lost, and hence the individual is operating at a reduced motivation. The requirement to attend carefully to the material in order to get a correct response is dissipated where the material itself is simple, and hence the strength of a self-reinforcement would be dissipated, along with the attention factor. The quality of the self-reinforcement, therefore, is dependent upon some level of difficulty. This is further substantiated by the fact that the ninth graders, when dealing with the Type 6 noun determiners where the probability of a correct response is .33, did relatively less well than the seventh graders. We can assume that the ninth graders would find this material exceptionally easy and therefore without motivating value which would overcome the amount of energy output required to read the material and answer the cloze responses. The seventh graders, on the other hand, find this level of difficulty to be highly motivating. This is confirmed not only in the present experiment but also in the preceding pilot studies. These data confirm the hypothesis that it is possible to have comprehension materials which are too simple for a particular group to achieve maximum comprehension.

Type Comparisons. In general, at all grades, Type 2 (noun deletions), Type 4 (adjective deletions), and Type 3 (verb deletions) tend to be inferior to the control group, Type 0, and to Type 1 (random deletions), Type 5 (preposition and conjunction deletions, and Type 6 (noun determiner deletions).

Selections containing Type 0, Type 5 and Type 6 also contain the greatest amount of information and present the least difficulty for the students.

Recall of Comprehension Materials. The recall test was comprised of thirty-six questions. Six of the twelve comprehension questions which followed the selections were randomly chosen for the comprehension test. Each student was assigned a form of the test commensurate with the selections which he had previously read. The recall test was given ten weeks following the last of the selections. The means and variances are presented in Table XIV. The rank order for recall comprehension indicates that Type 6 (noun determiners) is first, followed by Type 5 (prepositions and conjunctions), Type 2 (nouns), Type 3 (verbs), Type 4 (adjectives), Type 0 (no deletions) and Type 1

Table XIV

RECALL COMPREHENSION TEST SCORE BY DELETION TYPE

	Deletion Type						
	0 Control	1 Random	2 Noun	3 Verb	4 Modifier	5 Prep.& Conj.	6 Noun Determ.
$\bar{X}$	2.56	2.44	2.85	2.65	2.58	2.92	2.89
N of Tests	284	171	194	106	163	303	292
S^2	3.92	1.44	1.16	1.23	1.10	1.67	1.70

(random deletions). Types 5, 6, and 2 were each found to be significantly superior to the Type 1 random deletions for recall comprehension: t 2,1 = 2.10; t 6,1 = 2.97; and t 7,1 = 2.30. No other significant differences were found.

It should be noted here that the Type 5 and Type 6 were, for all grade levels, typically among the highest scores on the initial comprehension test, whereas Type 2 was quite typically one of the lowest. Furthermore, the Type 0 comprehension exercise material was again among the highest, and on the recall was among the lowest. This raises a question as to the relationship between the immediate learning and the recall of information. Results from learning experiments show that the greater the initial learning, the greater the recall. However, under present experimental conditions each subject was responding to a single exposure. Following the reasoning in the field that recall is related to initial learning, one would expect relatively high correlations between immediate comprehension test scores and recall scores. This proves not to be the case in the present instance. The correlations range from a +.76 to a -.29 with an average determined by z-transformation of .20. Further, it would seem highly probable that the number of correct closures should relate to immediate comprehension test scores. These range from .76 to -.11 with an average determined by z-transformations of .20. Further, there appears to be no relationship between the number of correct closures and the number of recall comprehension questions answered. Correlations here range from +.56 to -.92, with an average of -.019. The indication here is clear. Cloze procedure itself seems to be testing something different from the usual type of comprehension question. This result would support other findings such as those by Weaver and Kingston[46] in the same area. Further, initial reading comprehension test scores, whether cloze procedure of not, are testing something different from final recall procedures, which would indicate perhaps that if one is concerned with the long-term remembering of materials, the procedure of paragraph comprehension exercises which we customarily use to teach reading comprehension is not adequate. In essence the long-term recall of these materials appears to be a quite different skill from the reading comprehension skills which are taught via paragraph comprehension techniques.

Sex Comparisons by Grade. Tables XV and XVI show the means and variances for boys and girls by grade level. An analysis of variance indicated significant differences between grades but no significant difference between the sexes. The data show that in thirteen instances out of twenty-eight, the boys achieved superior mean scores. This difference was significant in two instances. The comparisons between boys and girls at the ninth grade level shows the boys to be significantly superior (Type 0, t = 3.98, Type 1, t = 3.69). The girls appeared to be superior in fewer instances as the grade level advanced. Fifth grade girls were superior to the boys in three instances: Type 1, random deletions (t = 2.06); Type 3, verbs (t = 3.15); and Type 6, noun determiners (t = 2.65). At the seventh grade level girls were significantly superior to boys for Type 2, nouns and Type 6, noun determiners: (t = 1.96 and 2.70 respectively). At the ninth grade level the girls were superior to the boys in one instance, noun determiners (t = 3.50); and at the eleventh grade level girls were superior to boys in one instance,

Table XV

MEAN CORRECT RESPONSES TO COMPREHENSION QUESTIONS

BY GRADE LEVELS AND TYPES OF DELETIONS FOR BOY SUBJECTS

		Deletion and Type						
		0	1	2	3	4	5	6
Grade Level	Statistic	Control	Random	Noun	Verb	Modifiers	Prep & Conj	Noun Determ
5th	N	136	81	94	96	75	102	121
	$\bar{X}$	5.53	4.43	5.09	4.68	4.47	5.85	5.45
	S^2	5.35	4.33	5.44	5.46	5.69	5.99	6.08
7th	N	99	97	65	42	67	99	96
	$\bar{X}$	6.80	6.26	4.74	5.64	5.78	6.86	7.20
	S^2	6.76	5.78	18.64	5.80	5.91	6.47	6.87
9th	N	93	98	84	94	64	116	94
	$\bar{X}$	8.84	7.91	7.00	7.09	6.77	7.87	6.22
	S^2	3.46	7.71	8.63	5.37	5.40	5.47	6.71
11th	N	111	78	59	66	76	16	43
	$\bar{X}$	8.37	6.85	6.71	7.53	6.53	7.81	7.74
	S^2	5.02	4.43	5.17	7.26	6.33	3.07	15.56

Table XV (continued)

MEAN CORRECT RESPONSES TO COMPREHENSION QUESTIONS

BY GRADE LEVELS AND TYPES OF DELETION - GIRL SUBJECTS

Grade Level	Statistic	0 Control	1 Random	2 Noun	3 Verb	4 Modifier	5 Prep & Conj	6 Noun Det.
5th	N	64	102	136	94	33	82	59
	$\bar{X}$	5.97	5.13	5.21	5.74	4.79	5.48	6.42
	s^2	4.13	6.40	4.97	5.53	4.81	4.40	4.98
7th	N	90	59	59	67	114	73	64
	$\bar{X}$	6.40	6.20	5.88	6.28	5.46	6.92	8.17
	s^2	5.37	5.66	3.07	5.03	7.17	4.31	3.67
9th	N	136	44	43	80	119	61	83
	$\bar{X}$	7.75	6.05	6.53	7.06	7.34	8.18	7.53
	s^2	5.21	7.72	6.07	8.37	4.20	5.82	5.70
11th	N	96	93	92	74	65	96	76
	$\bar{X}$	8.05	7.15	6.92	7.39	7.26	7.64	7.66
	s^2	4.60	6.92	7.37	3.44	6.05	5.49	5.43

random cloze (t = 4.70). Data suggest that girls, therefore, appear to be more responsive to the noun determiners than do boys, in that in three instances out of four they were significantly superior; and the data also suggest that the difference between boys and girls, which may be present in the early years, seems to disappear as the children get older.

Girls at the seventh and ninth grade levels seem to be better able to reinforce themselves from their language patterns than are boys. A comparison between the grades shows that ninth grade boys are superior to eleventh grade boys in Type 1 (t = 2.98) and inferior to eleventh grade boys in Type 6 comprehension (t = 2.36). The mean score for boys in Grade 9 is superior to the mean score for boys in Grade 11 in five of the six comparisons. This may be a function of the differential approach to language materials of boys and girls, and indicate, as do the other data, the tendency for younger boys to be self-reinforced by correct responses in close procedure materials is somewhat less than that of girls.

DISCUSSION

The data confirmed the hypothesis that essentially higher grade pupils confronted with material of extreme simplicity, approximately six reading levels below their mean ability, could not derive sufficient self-reinforcement from reading the material or completing the cloze exercises to comprehend at a level better than that of the ninth grade.

The sex comparisons indicate that, as the material gets relatively more simple, the differences between the sexes seems to disappear and the tendency for boys' performance to be superior over girls increases. This suggests testing the hypothesis that teaching boys at lower levels with material which is somewhat simpler may erase the traditional difference between boys and girls in reading ability. Type 0, 5 and 6 materials confirm the hypothesis of self-reinforcement. It remains, however, to determine why the other deletion types do not follow this pattern. The clue to this dilemma is found in the decreased performance of the ninth graders on Type 6 (noun determiners).

The fact that relative comprehension decreases with simple material gives rise to a requirement to examine school materials, particularly at the first grade level, to determine if the simplicity at that level is not debilitating to reading progress.

The data also suggest extreme care with the pre-reading survey techniques. The effect of the pre-reading survey may be to reduce the value of the reading material in the eyes of the individual such that he will not then comprehend the material. Obviously this caution would not hold where the material is new to the reader. However where the material is familiar, pre-reading may cause him to pre-judge and to underestimate the value of the material. Under these conditions survey techniques will reduce rather than increase reading comprehension.

The experiments suggest caution in the interpretation of certain test results. It is common practice in the building of standardized tests to extrapolate the performances of the extremes of the standardization group from data found at the middle of the standardization range. This practice may have the net effect of making completely unreliable scores at the upper extremes of the range. At the upper limits of a standardization group where self-reinforcement is relatively low, reading test scores would be more dependent upon motivation than upon actual reading ability. Working at the extremes of the range for reading tests standardized in the above manner, pre-test and post-test differences would tend to be indicated less often than they occur. Further, a standardized comparison of one group with another would not reflect a true index of differences between the groups. Ranking individuals within the class would be more likely to rest on motivational variables rather than upon reading ability. This may account for the fact that many studies of college reading programs have not found significant reading comprehension differences. The data lend credence to the test-makers admonition that one should select a test where the group to be tested is close to the middle of the range of the scores.

The data also suggest further exploration of the notion that cloze procedure is measuring a comprehension skill which is different from the traditional paragraph, question technique. The highly variable correlations between cloze procedure and reading comprehension indicate that serious exploration may reveal new and useful dimensions for the measurement of reading skills.

CHAPTER VII

THE EFFECTS OF NON-REINFORCED CLOZE PROCEDURE UPON THE ABILITY TO MAKE CLOSURES

The previous chapter found a differential in the ability to make closures based upon the syntax, and indicated a self-reinforcing factor in operation during a short sequence of cloze exercises. The duration of a long series of cloze exercises where no overt reinforcement is present takes on the aspects of extinction trials. We can expect therefore some effects in terms of the ability to make closures during a prolonged sequence of non-reinforced cloze exercises.

In an earlier study, Bloomer and Heitzman found an inequality in the ability of individuals to make closures of varying syntaxes. The question arises, particularly when dealing with self-reinforcement, as to whether this difference in difficulty will in turn affect variables such as the ability to make closures, and reading comprehension, and more particularly will practice with cloze procedure materials of a particular syntax increase the ability of an individual to make closures of that same syntax.

With the postulation in the previous chapters that an individual is able to reinforce himself by producing language patterns and matching those with the language patterns which he has already learned, one would expect some increase in the ability to perform cloze material subsequent to practice with non-reinforced cloze exercises. Comparisons, therefore, between the control group (that is, those persons who performed the exercises with no closures) and cloze procedure experimental groups who performed cloze procedures of various syntactical deletions should indicate a superiority on the part of the experimental groups to perform cloze procedure materials.

Here again we would expect that some of the syntact deletion types would not generate sufficient feedback for individuals to learn. Others where the probability of reinforcement is sufficiently high would not produce sufficient reinforcement for an individual to make closures. Therefore one would hypothesize that the relatively more difficult cloze procedure exercises and the relatively more simple cloze procedure exercises would not produce an increase in the ability to make closures. However, those of moderate difficulty should increase the ability of individuals to make closures.

The effects of extinction, however, should result in a shift in perceived difficulty such that all materials should appear more difficult. Therefore under conditions of the present experiment one would hypothesize that if an increase in the ability to make closures occurs it will be most likely found in the easier samples (i.e., function words).

With the postulation in the previous chapters that an individual is able to reinforce himself by producing language patterns and matching those with the language patterns which he has already learned, one would expect some increase in the ability to perform cloze material subsequent to practice with non-reinforced cloze exercises. Comparisons, therefore, between the control group (those persons who performed the exercises with no closures) and cloze procedure experimental groups who performed cloze procedures of various syntactical deletions should indicate a superiority on the part of the experimental groups to perform cloze procedure materials.

Here again we would expect that some of the syntact deletion types would not generate sufficient feedback for individuals to learn. Others where the probability of reinforcement is sufficiently high would not produce sufficient reinforcement for an individual to make closures. Therefore one would hypothesize that the relatively more difficult cloze procedure exercises and the relatively more simple cloze procedure exercises would not produce an increase in the ability to make closures. However, those of moderate difficulty should increase the ability of individuals to make closures.

PROCEDURE

Materials. The materials for the present study consisted of twenty-four short articles six hundred words in length. The articles were selected such that a complete story was told. Material was all fiction. Twelve comprehension questions were written for each of the articles. The twelve comprehension questions were pretested to discriminate between the top and the bottom students in the pre-test group. The reading level of the articles was between 5.0 and 6.0 on the Yoakum formula. The articles were then adapted to cloze procedure in the following manner:

Type 0. The Control Group; zero deletions
Type 1. Ten percent random deletions:
Type 2. Ten percent noun deletions
Type 3. Ten percent verb deletions
Type 4. Ten percent modifier deletions
Type 5. Ten percent preposition and conjunction deletions
Type 6. Ten percent noun determiner deletions.

Deletions were done in the usual way. The initial word to be deleted was selected at random and every tenth word thereafter was deleted. Where specific syntax was involved the nearest word of the appropriate syntax to the tenth word was deleted. Other grammatical categories were omitted since they did not include ten percent of the running words.

Sample. The sample consisted of 149 ninth grade pupils of a suburban central school in up-state New York. The pupils were randomly assigned to deletion type. Each pupil received exercises of only a single deletion type. The material was worked in class on alternate days, two exercises per week, for twelve weeks. All experimental treatment groups were to be found in each class to control for class assignment variables. The students were assigned to four different orders of presentation of the exercises to control for order of presentation variables. The students wrote the correct responses to the closures on a separate answer sheet. The subsequent reading comprehension questions were marked on IBM sense cards for processing directly through the computer. There were no significant differences between various orders of presentation, hence all orders of presentation were treated together. There were no significant differences between boys and girls, hence sexes were treated together.

The pupils were given a pre-test. Subsequent to the experiment the pupils were given a final cloze test consisting of a thousand-word passage with twenty words deleted from each of the five syntactic types listed above. This test allowed the comparison of the ability of the various groups to make closures, as well as the ability to make closures of a specific syntactic type. The pupils were also given the California Test of Mental Maturity as a part of the post-test battery.

The Test. A cloze procedure test was administered to the groups one and one-half weeks subsequent to the termination of the experiment. The test itself was a cloze procedure exercise one thousand words in length with a ten percent random deletion. Twenty of the deletions were of each of the five syntactic types used in the study; that is: nouns, verbs, modifiers, prepositions and conjunctions, and noun determiners. The syntax of the deletion types was randomly ordered within the test, and where the tenth word was not of the assigned deletion type, the nearest word of the appropriate syntax was deleted. The subjects were required to fill in the missing words. The tests were scored for the number of correct responses at each deletion type as well as the total number of correct responses. In the present study the zero deletion type served as a control group, having read and answered the comprehension questions for all the material but having had no practice in cloze procedure.

RESULTS

Closures by Exercise. Table XVI gives the mean number of correct closures for Exercises 1, 10 and 22 by deletion type. Different numbers of correct responses were made for the varying deletion types. These seem to fall into three groups. Random deletions, noun deletions, and verb deletions were approximately equal. On the other hand, very few of the modifier deletions were answered correctly, while both prepositions and conjunctions, and noun determiners, the function words, tended to be answered correctly by a relatively large proportion of the students. The number of correct closures made on Exercise 10 was smaller than the number of closures for Exercise 1, excepting the noun determiners. Similarly, for Exercise 22 the number of correct closures for all categories except modifiers and noun determiners was smaller than the number of correct closures for Exercise 1, indicating that in general the individuals tended to do less well as the sequence with non-overt reinforced cloze procedure exercises continued. Table XVII indicates this tendency was not significant for the group receiving random deletions. There were no significant differences for modifiers, or

Table XVI

MEAN NUMBER OF CORRECT CLOSURES FOR EXERCISES 1, 10 AND 22*

BY DELETION TYPE

Deletion Type		Exercise					
		1		10		22	
	N	X̄	SD	X̄	SD	X̄	SD
2 Random	14	25.94	5.32	20.13	9.96	24.73	7.34
3 Noun	12	22.36	6.57	13.92	9.66	13.90	5.43
4 Verb	11	22.67	8.01	14.56	8.04	14.00	9.34
5 Modifier	15	14.26	3.66	13.00	6.85	15.66	6.49
6 Preposition and Conjunction	25	37.56	6.62	36.88	8.17	35.52	8.49
7 Noun Determiners	21	37.10	11.13	47.05	5.93	41.44	13.49

*(Sixty possible)

Table XVII

't' TEST COMPARISON OF NUMBER OF CORRECT CLOSURES

IN EXERCISE 1 WITH EXERCISES 10 AND 22

Deletion Type	Comparison Exercise	
	10	22
1 Random	NS	NS
2 Nouns	3.62	4.97
3 Verbs	2.37	2.32
4 Modifiers	NS	NS
5 Prepositions and Conjunctions	NS	NS
6 Noun Determiners	3.63	NS

prepositions and conjunctions, and noun determiners were reproduced significantly better in Exercise 10. However, this same condition did not hold for Exercise 22.

The data show the differential effects of rate of self-reinforcement upon extinction. The random material, which has variable difficulty and thus approaches random internal reinforcement, does not suffer the effects of extinction. The function words where the rate of reinforcement is relatively higher similarly show no extinction effects. The noun determiners, where the actual rate of reinforcement is relatively high and because most of the words are interchangeable giving a still higher rate of self-reinforcement, shows an increase similar to that found in the pilot studies with this same deletion category with seventh grade pupils. The increase appears to be affected by the extinction process in Exercise 22.

In interpreting these data it should be cautioned that the subjects for whom data is presented are a select group. All pupils who had not completed at least eighteen exercises were deleted from the population for analysis. The group which was not treated are presumably those children who suffered the effects of extinction to the greatest extent. It should be noted that the N for each deletion type in the present study, as compared with the original N given in Chapter II, reflects in all instances a marked drop in number of subjects and reflects the relative difficulty of the deletion types.

The lack of a significant drop in the modifier deletion group may be seen as a function of this selection process, as well as the fact that a relatively wide variety of incorrect modifiers will still be acceptable and hence self-reinforcing. While the specific modifier is difficult to fill in, an individual subject is more likely to fill in one which satisfies all the syntactic clues than is the case with nouns or verbs. Hence the rate of self-reinforcement, as with noun determiners, is greater than the objective rate indicated by the scoring.

Closures by Deletion Type. Tables XVIII, XIX, and XX give the significant 't' rations between deletion types for Exercises 1, 10 and 22 respectively.

For Exercise 1, the random deletion group was significantly superior to nouns, Verbs and modifiers in the number of correct deletions. Deletion Type 5 (prepositions and conjunctions) and Deletion Type 6 (noun determiners) were superior to all other types in the number of correct closures.

For both exercises 10 and 22, the deletion types 5 and 6 are significantly superior to all other types. Random deletions (Type 1) shows the effects of limited extinction in that it is not significantly superior to noun or verb deletions.

The final cloze test was administered one and one-half weeks subsequent to the last cloze exercise, which allowed time for any spontaneous recovery to occur. The No Deletion group was included in this testing to act as a control comparison group for those trained upon cloze procedure materials.

Table XXI gives the mean and standard deviation of the correct responses for the final cloze test. The mean number of correct

Table XVIII

SIGNIFICANT 't' RATIOS BETWEEN NUMBER OF CORRECT CLOSURES FOR EXERCISE 1 BY DELETION TYPE

Deletion Type			Deletion Type				
	N	1	2	3	4	5	6
1			3.68	2.81	3.12	-3.81	-4.31
2						-8.77	-7.62
3						-6.01	-6.17
4						-7.70	-6.98
5							

Table XIX

SIGNIFICANT 't' RATIOS BETWEEN NUMBER OF CORRECT CLOSURES FOR EXERCISE 10 BY DELETION TYPE

Deletion Type			Deletion Type				
	N	1	2	3	4	5	6
1					2.24	-5.89	-9.54
2						-7.07	-10.73
3						-7.07	-10.86
4						-9.23	-14.30
5							- 4.78

Table XX

SIGNIFICANT 't' RATIOS BETWEEN NUMBER OF CORRECT CLOSURES FOR EXERCISE 22 BY DELETION TYPE

Deletion Type			Deletion Type				
	N	1	2	3	4	5	6
1					7.15	-6.19	-4.03
2					3.69	-6.38	-4.70
3					2.97	-5.00	-4.00
4						-14.32	-8.76
5							

Table XXI

MEAN FINAL CLOZE SCORES FOR TEST DELETION BY STUDENTS

COMPLETING EACH TYPE OF EXERCISE DELETION

Final Cloze Deletion Type	Statistic	Exercise Deletion Type 0 Control	 1 Random	 2 Noun	 3 Verb	 4 Modifier	 5 Prep & Conj	 6 Noun Det.
3. Nouns	$\bar{X}$	3.65	4.71	4.22	3.63	3.35	4.94	4.44
	SD	1.89	2.39	3.22	1.97	1.46	2.79	3.12
4. Verbs	$\bar{X}$	3.09	3.92	3.12	5.13	4.23	4.31	3.89
	SD	1.41	1.75	1.93	2.58	1.65	1.85	2.44
5. Modifiers	$\bar{X}$	2.85	3.57	2.56	3.88	3.54	3.76	3.14
	SD	1.72	1.71	1.85	2.17	1.68	1.61	1.64
6. Prepositions and Conjunctions	$\bar{X}$	7.43	10.00	7.78	8.62	9.69	11.53	9.00
	SD	3.49	3.75	5.08	4.53	3.05	4.25	5.24
7. Noun Determiners	$\bar{X}$	9.61	14.93	8.44	13.75	13.31	12.47	13.94
	SD	4.93	4.57	5.84	5.23	2.63	4.19	4.72
Total Cloze Deletions	$\bar{X}$	26.17	35.56	27.56	35.00	34.62	36.38	33.21
	SD	11.42	12.71	16.06	14.15	7.46	13.69	14.80

responses for the experimental groups were numerically greater than the control group in all but four instances. Two of these instances occurred in the group which had noun deletions, and one each for verbs and modifiers. The comparison of the total number of responses for the experimental and control groups again shows in each instance the number of correct responses in the experimental group was numerically greater than for the control group.

't' Test Comparisons

Experimental Group 1. Experimental Group 1 which was trained with random closures showed a significant superiority to the control group in total number of responses. The contribution to the superiority lay primarily in Deletion Type 6, noun determiners, and Deletion Type 5, prepositions and conjunctions, indicating that the major learning of individuals who performed non-reinforced random closures is in the area of the function words.

Experimental Group 2. For pupils trained on noun deletions, no significant differences occurred between that group and the control group.

Experimental Group 3. The group that was trained on verb deletions was significantly superior to the control group, as would be anticipated from the hypothesis. This same group approached significance, t = 1.96, for the number of noun determiners correctly answered in the test. However, there were no other significant comparisons.

Experimental Group 4. The experimental group who practiced with modifier deletions performed significantly better for total number of responses on the final test, as well as for Type 6 (noun determiners)

Table XXII

't' TEST COMPARISONS OF CONTROL GROUP (ZERO DELETIONS) WITH TYPE OF EXERCISE DELETION BY TEST DELETION TYPE

Test Deletion Type	Exercise Deletion Type 1 Random	2 Noun	3 Verb	4 Modifier	5 Prep & Conj	6 Noun Det.
2. Nouns						
3. Verbs			2.10	2.06	2.33	
4. Modifiers						
5. Prepositions and Conjunctions	2.16			2.03	3.36	
6. Noun Determiners	3.46		1.96	2.93	2.04	2.86
Total	2.36			2.67	2.58	

and Type 5 (prepositions and conjunctions), and Type 3 (verbs). Here again the ability to perform better with function words contributes to the major portion of the superiority of this group, but included also is the ability of this group to fill in missing verbs.

Experimental Group 5. The group which had been trained on prepositions and conjunctions was significantly superior to the control group for the total number of responses for Type 6 (noun determiners); Type 5 (prepositions and conjunctions) and for Type 3 (verbs). In addition, this group approached significance (t = 1.67) for Type 4 (modifier deletions) and for Type 2, noun deletions (t = 1.65). If a one-tailed test were used this group would be significantly superior to the control group for modifier deletions.

Experimental Group 6. Experimental Group 6, trained on noun determiner deletions, had the highest probability of correct response, as would be expected by the self-reinforcement hypothesis. This group did not do as well as those groups with the moderately difficult deletion types. Experimental Group 6 approached significance in the total number of correct responses (t = 1.70) and with the one-tailed test this would be significant beyond the .05 level of confidence.

The data indicate that whether or not the reading of an individual is affected, his ability to fill in materials of differing syntax forms can be improved by the use of non-reinforced cloze materials of modifier or preposition and conjunction deletion types.

The use of cloze procedure, particularly of two varieties, may prove to be of considerable significance in teaching individuals grammar and syntax rules, albeit these rules represent a relatively small percentage of the number of total areas. The effect of cloze procedure is therefore to teach the individual to fill in the specific correct word rather than a word of appropriate syntax but relatively incorrect. In this sense the use of non-overt cloze procedure tends to conventionalize the language of the individual. Using non-reinforced cloze procedure and comprehension questions related to it, individuals given a sustained series tend to do somewhat worse as time progresses. One would expect that this is fundamentally the result of extinction with the cloze procedure itself as there is no apparent effect of the prolonged series of cloze procedure upon the comprehension of individuals. In essence individuals therefore tend to get saturated with non-reinforced cloze procedure and this saturation appears to show up before ten of these have been completed. The question remains whether or not a similar situation occurs with reinforced cloze procedure. Studies with reinforced cloze procedure and motivation have indicated that such a saturation does not occur and that different results are obtained in terms of the reading comprehension of individuals who work with random cloze material in a graded sequence; they do get better at reading.

DISCUSSION

The results of the study suggest that there is a differential in the capacity of cloze procedure exercises of varying syntactic deletions to produce an increase in the ability to perform cloze procedure material.

Verb deletions or noun determiner deletions both act rather specifically to increase the ability of an individual to fill in verbs or noun determiners, respectively. When one considers modifier deletions or preposition and conjunction deletions, the ability to make closures apparently generalizes to other syntactic types. This is particularly true for preposition and conjunction deletions which indicate use of the non-reinforced technique the most effective way to teach closures.

It is also significant to note that for Type 2 deletions, no increase in the ability to make closures occurs. This would indicate that filling in noun deletions may involve a somewhat different process. Typically when one is doing cloze procedures with noun deletions, one may merely substitute for a noun blank the concept thing, place, or object in general rather than a specific. The meaning of the material is not necessarily therefore contingent upon the specific noun name but is relatively clear when pronouns or merely one of the noun concepts is substituted. Thus the individual who has nouns to fill in has fundamentally a ready substitute set of concepts and does not necessarily have to attend sharply to the meaning in order to figure out what that meaning would be. On the other hand, the syntactic rules for a specific noun or verb are more restrictive in the cloze situation than for other parts of speech. The individual given verbal cloze procedure material still has a generic notion of action as a possibility for filling in the verb. However, the actions tend to be more specific and to relate much more to the meaning of the paragraph than do the nouns. The verb is limited somewhat by the noun as well as by the structure of the rest of the material. One must therefore be cognizant of the possible verb action of any noun in order to make sense of the cloze procedure material. A certain larger amount of focus is therefore required if an individual is to fill in verbs. He must be cognizant at least of the noun about which he is talking. Modifiers are the most difficult to fill in simply because a larger range of modifiers will fit any particular modification situation and there is no broad overall concept which may be substituted. While modifiers do not necessarily carry meanings in a cause and effect situation they none-the-less carry a large proportion of the more subtle meaning of the language and are the words which tend to differentiate a particular passage from others of a similar nature. To fill in a modifier, therefore, an individual must be cognizant of the nouns and of the verbs, and then must choose from a relatively large number of modifiers the appropriate modifier to fit the particular noun-verb relationship as well as the intention of the meaning of the paragraph.

Much the same is true of prepositions and conjunctions. They tend also to function to modify the meaning of the material and hence require more focus on the part of the pupil than do some others. It is true that prepositions and conjunctions in some measure involve generic concepts but in each instance there are typically several of these which may be chosen, and it is sometimes difficult to determine which one would be correctly applied to any particular situation. Prepositions and conjunctions also tend to make relatively simple closures in that the probability of deriving a correct closure is relatively high.

The noun determiners typically seem to be of relatively little importance as their inclusion or exclusion affects the meaning of the material to a relatively small degree. Typically again the probability of a correct response is extremely high.

At the present time with the emphasis of linguists upon the fact that children have developed syntactic sequences prior to school, the problem the school faces is often one of changing these syntactic patterns to a more standard form. Children, even those from homes of deviant linguistic patterns, tend to be aware of standard patterns through outside contacts and mass media. Standard speech is thus a function primarily of differential reinforcement of the standard over the deviant pattern. Since most individuals develop several linguistic patterns to function in the various roles they play, the school's concern is with ensuring that the more formal standard patterns are available to the individual for situations where these patterns are advantageous to the individual. The cloze procedure may have some utility in the teaching of syntactic patterns. It presents a series of language problems in sequence which would help the child to learn to develop language patterns which, at least for himself, are self-reinforcing. Cloze procedure therefore would allow a teacher the technique for teaching the child to produce satisfactory language sequences in a relatively controlled situation. The use of the concept of fading in a fashion similar to that described by Skinner, where one would perhaps increase the percentage of deletions required in material, would teach the individual to produce longer and longer appropriate language sequences, whereas the initial material would still offer a framework in which this work could be done. Thus the children would be taught to fill in words of appropriate syntax and meaning in relatively logical sequences. It would appear that it would be superior to use reinforced cloze procedure materials rather than non-reinforced cloze procedure as the effectiveness of reinforcement in shaping would then be taken advantage of. Furthermore, an individual should function at some specific level of difficulty of closure or syntactical deletion until some criterion were reached, at which point the individual would move to more difficult material. It is also clear that grading should not be done for the specific word which the author included, but rather that grading should be done on the basis of words of appropriate meaning, with preference given for those words which are unusual or relatively less frequent. This would be more and more true as the percentage of deletion increased and the skill of the pupil became greater. The present experiment would suggest that the use of preposition and conjunction deletions at least at the outset of the teaching would be superior. However, it must be remembered that this is a function of the non-reinforced situation in which the material was presented. In a reinforced situation some other form or forms of deletion would appear to be superior. Certainly if a child were having difficulty with words of a particular syntactic type, cloze procedure materials, excluding words of this particular type, would serve as a training device for this particular child. In any case, before using a deletion type such as prepositions and conjunctions, an experiment should be performed to determine whether the effectiveness of prepositions and conjunctions as found in the present experiment is a function of either the degree or nature of self-reinforcement involved, or whether it is a function of the nature of the preposition and conjunction. If the former is the case we may find using reinforced material less effective than some other form of deletion. If the latter is the case it would suggest that the filling in of prepositions and conjunctions would be superior for at least the initial learning for such a teaching of language sequences.

SUMMARY AND CONCLUSIONS

A series of twenty-four cloze procedure exercises followed by twelve comprehension questions was given to 123 ninth grade pupils. Each individual was randomly assigned to a particular kind of deletion ranging from no deletions as the control group to random deletions, noun deletions, verb deletions, modifier deletions, preposition and conjunction deletions, and noun determiner deletions. All deletions were at the ten percent level. Data for the first exercise, the tenth exercise and the twenty-second exercise were presented. The results indicate the following:

1) With the exception of modifier deletions and noun determiner deletions, individuals tend to make fewer correct closures as the process continues.

2) Individuals tend to make more correct noun closures in the tenth exercise.

3) Individuals who received random deletions, modifier deletions, or preposition and conjunction deletions tended to do better at filling in correct closures than did the control group with no deletions, on

the final test.

4) Where cloze procedure is used as a part of a sequence of tests, one can expect a decrease in score as a function of the repetition of cloze material.

5) Cloze procedure should have utility as a technique for teaching standard syntactic patterns to children of substandard language.

CHAPTER VIII

THE EFFECT OF NONOVERT REINFORCED CLOZE PROCEDURE UPON READING COMPREHENSION

Reinforced cloze procedure has been used in a form somewhat different from the present experiment as a teaching device for reading comprehension. The study of Bloomer (1962) shows that where cloze procedure is reinforced in a motivational scheme, effective increases in reading comprehension occur. There appear to be two strands of research dealing with the cloze procedure. The first, using cloze procedure as a readability or reading-testing technique, uses non-reinforced cloze procedure. The second uses cloze procedure as a teaching device, employs some form of knowledge of results or reinforcement subsequent to the closures, as we saw in Chapter VI. It can be hypothesized, however, that non-reinforced cloze procedure has built into it a certain element of reinforcement in that, upon completing a closure, the individual is able to reinforce himself by matching his language patterns with those of the author, or rather those he has reconstructed from his own and the author's patterns. This would lead one to expect that the non-overt reinforced cloze procedure might also have involved certain elements of teaching which would tend toward improvement as an individual proceeded through a series of cloze tests. The present study therefore is an attempt to measure the effects of a protracted series of non-reinforced cloze procedure exercises upon the reading comprehension ability.

Two methods of measuring the effects of non-overt reinforced cloze procedure upon reading comprehension are available in the present experimental design. The first is through the use of the comprehension tests subsequent to the cloze exercises. If non-overt reinforced cloze is effective as a method of teaching reading comprehension, one would expect an increase in immediate comprehension scores or in final comprehension scores, as the series continued.

The second measurement technique involves the use of standardized tests to measure pre- and post-test differences. Here significant differences between control and experimental post-tests would attest to the efficacy of non-overt reinforced cloze as a comprehension teaching tool.

PROCEDURE

Prior to the onset of the experiment, the Differential Aptitude Test had been given to the pupils as a pre-test. The post-test, the Iowa Test of Basic Skills, was administered to ninth grade pupils subsequent to the completion of the exercises. The sections administered were: the language section, the vocabulary section, and the reading section.

Pre- and post-test scores were obtained for 125 ninth grade pupils who had each completed at least eighteen cloze procedure exercises.

In addition, each of the cloze exercises taken by the students was in effect a reading comprehension test, with the dozen reading comprehension questions which followed it. For present purposes, the number of correct reading comprehension questions for the first, fourth, seventh tenth, fourteenth, sixteenth, nineteenth and twenty-second comprehension exercises were also examined. If the non-reinforced cloze procedure were to have an effect upon reading comprehension one would expect an increase in correct responses toward the end of the sequence.

RESULTS

Table XXIII gives the pre-test. No significant differences were found for the DAT verbal scores. One difference approached significance, the difference between Type 0 (the control group) and Type 5 (prepositions and conjunctions): 't' was equal to 1.97 with 44 degrees of freedom. Essentially, then, the randomly selected groups were approximately equal in verbal ability as measured by the Differential Aptitude Test at the outset of the experiment.

Table XXIV gives the means and standard deviations for the post-test scores from the Iowa Test of Basic Skills. The significant differences of the post-test involved Type 4 (modifiers) over the control group, and Type 5 (prepositions and conjunctions) over the control group for the spelling sub-test (L-1) of the Iowa Test of Basic Skills; the comparison of the modifier deletion group with the control group (t = 3.14); and the comparison of the preposition and conjunction deletion group with the control group (t = 2.67). No other significant differences are found. However, Type 6 approaches significance over the control group for the vocabulary sub-test where t = 1.97 with 47 degrees of freedom. The data do not indicate, therefore, that the use of non-reinforced cloze procedure material produces significant increase in reading or in language skills except in spelling, where the deletion of either modifiers

Table XXIII

PRE-TEST DIFFERENTIAL APTITUDE TEST MEAN SCORES

D.A.T. Subtest

Deletion Type	N	Statistic	(9) Verbal Reasoning	(10) Arithmetic Reasoning	(11) Spatial Reasoning	(12) Mech. Reasoning	(13) Clerical	(14) L-SP	(15) L-SE
0. No Deletion	27	$\bar{X}$	6.37	6.87	6.90	5.61	5.07	5.08	5.29
		SD	2.25	2.35	1.81	3.05	2.75	3.43	2.55
1. Random	15	$\bar{X}$	7.25	6.37	5.43	5.73	4.77	5.57	6.37
		SD	2.30	2.03	2.59	3.34	2.52	2.94	2.57
2. Nouns	11	$\bar{X}$	6.78	6.83	6.43	6.50	6.56	4.89	5.50
		SD	3.26	2.72	2.69	3.33	2.68	2.75	2.33
3. Verbs		$\bar{X}$	6.64	6.14	6.01	6.42	5.36	5.57	5.93
		SD	1.86	3.58	3.45	3.31	3.02	2.30	3.22
4. Modifiers		$\bar{X}$	7.32	7.47	8.15	6.80	4.79	6.53	5.75
		SD	1.87	2.74	1.28	3.51	1.50	2.32	2.23
5. Prepositions and Conjunctions		$\bar{X}$	7.60	6.78	7.10	6.35	6.02	7.37	7.03
		SD	1.97	2.60	2.54	2.69	1.96	2.19	2.31
6. Noun Determiners		$\bar{X}$	7.01	7.59	6.64	5.99	4.94	5.17	5.61
		SD	2.56	2.30	2.89	2.87	2.33	2.62	2.17

Table XXIV

POST-TEST. THE IOWA TEST OF BASIC SKILLS MEAN SCORES

			Iowa Subtest					
Deletion Type	N	Statistic	(27) L 1	(28) L 2	(29) L 3	(30) L 4	(31) R	(32) V
0. No Deletions	27	$\overline{X}$	9.50	9.83	10.11	9.35	9.68	9.93
		SD	1.64	1.51	1.73	2.02	1.38	1.46
1. Random	15	$\overline{X}$	10.13	9.13	9.54	9.52	10.04	9.90
		SD	1.75	1.72	1.34	1.00	1.25	1.66
2. Nouns	11	$\overline{X}$	9.91	9.53	9.82	9.79	10.09	10.43
		SD	1.22	1.53	1.88	1.64	1.41	1.19
3. Verbs	10	$\overline{X}$	9.39	9.12	9.37	9.31	9.92	9.83
		SD	2.67	2.60	2.75	2.31	2.12	2.31
4. Modifiers		$\overline{X}$	11.06	10.32	10.21	10.04	9.99	10.04
		SD	1.26	1.02	1.29	1.25	1.19	1.60
5. Prepositions and Conjunctions		$\overline{X}$	10.70	9.99	10.42	10.18	10.29	10.72
		SD	1.50	1.52	1.50	1.33	1.07	1.30
6. Noun Determiners		$\overline{X}$	10.30	10.04	9.86	9.87	10.34	10.21
		SD	1.46	1.74	1.76	1.65	1.32	1.40

or prepositions and conjunctions appears to increase spelling ability.

Individual Comprehension Tests Following Cloze Exercises. One would anticipate that if one learns through using non-overt reinforced cloze that the comprehension of passages would become somewhat superior with practice. On the other hand, since all of the information for comprehension is given to the control group with no deletions, one could further anticipate that this group would in turn become superior. No significant differences were found.

Table XXV gives the means and standard deviations for the number of correct comprehension questions for each type for exercises one, four, seven, ten, thirteen, sixteen, nineteen and twenty-two. Neither of the above suppositions appears to be the case.

Table XXVI gives the significant critical ratio for comparisons between the control group and the experimental groups. Significant differences indicating the superiority of the control group exist for exercise ten, where the control group (no deletions) is significantly superior to Type 1 (random deletions), Type 2 (noun deletions), and Type 4 (modifier deletions). On the other hand, for exercise thirteen, Type 5 (prepositions and conjunctions) is significantly superior to the control group. Again, in exercise twenty-two, the control group (no deletions) is superior to Type 1 (random deletions) and Type 6 (noun determiner deletions), and approaches significance ($t = 1.96$ with 28 degrees of freedom) for Type 3, (verb deletions). These data indicate clearly that the use of non-reinforced cloze procedure does not increase reading ability either during the process of practice with the cloze procedure or as a function

Table XXV

MEAN NUMBER OF CORRECT COMPREHENSION QUESTIONS BY EXERCISE AND DELETION TYPE

Deletion Type	N	Stat-istic	Exercise 1	4	7	10	13	16	19	22
0. No Deletions	27	$\bar{X}$	8.19	7.85	7.64	7.68	7.43	7.61	8.10	8.19
		SD	1.96	2.52	2.76	1.80	2.63	2.71	2.00	2.16
1. Random	15	$\bar{X}$	8.86	8.15	7.20	5.50	6.43	6.36	8.29	4.86
		SD	2.22	2.88	2.39	3.22	3.45	3.22	1.79	3.68
2. Nouns	11	$\bar{X}$	8.27	6.40	7.78	4.63	7.56	6.88	5.43	5.38
		SD	1.56	3.69	2.17	3.13	2.96	4.10	3.99	4.04
3. Verbs	11	$\bar{X}$	8.18	6.57	6.57	6.29	7.25	6.20	6.30	5.33
		SD	2.85	3.46	3.94	3.51	2.34	2.62	3.65	4.14
4. Modifiers	15	$\bar{X}$	8.80	8.58	7.73	5.56	8.17	7.31	8.10	6.13
		SD	3.01	2.16	2.49	2.86	1.53	3.33	1.53	3.10
5. Prepositions and Conjunctions	25	$\bar{X}$	8.76	8.50	8.10	7.40	9.60	8.44	8.33	6.89
		SD	1.67	2.50	2.28	3.09	1.50	2.65	2.41	3.69
6. Noun Determiners	21	$\bar{X}$	8.42	6.90	7.28	6.44	6.64	6.24	7.80	6.24
		SD	2.20	2.94	1.99	3.48	3.03	3.88	3.44	3.29

Table XXVI

SIGNIFICANT 't' TEST COMPARISONS OF CORRECT COMPREHENSION QUESTIONS FOR EXPERIMENTAL AND CONTROL GROUP (TYPE O) BY EXERCISE

Deletion Type	N	1	4	7	10	13	16	19	22
1. Random					2.16				2.40
2. Nouns					2.61				
3. Verbs									1.96
4. Modifiers					2.07				
5. Prepositions and Conjunctions						3.17			
6. Noun Determiners									2.11

of the post-test; nor do they substantiate the proposition that the greater amount of information present in the no-deletions material produces an increase in reading comprehension, though there were significant differences, most of which favored the Type O (no deletion) group. Nonetheless, there were many instances of negative comparison, although only one of these was significant. The present data indicate that non-reinforced cloze procedure material does not affect the reading ability of an individual any more or less than standard paragraph comprehension exercises. Further, it is indicated from the mean correct comprehension questions that neither non-reinforced cloze material nor practice with paragraph comprehension materials had any effect upon reading at all.

Correlations of Reading Test Scores. Table XXVII gives the correlation of the reading test scores with the number of correct closures for Exercises 1, 10, and 22, by deletion type. In general the correlations are positive, significant at the .01 level, and in the

Table XVII

CORRELATIONS OF READING TEST SCORES WITH NUMBER OF CORRECT CLOSURES FOR EXERCISES 1, 10 AND 22 BY DELETION TYPE *

Deletion Type	Pretest DAT VR			Post-Test Iowa Reading		
	1	10	22	1	10	22
1. Random	.34	.39	.65	.55	-.07	.65
2. Nouns	.75	.34	.45	.68	.27	.58
3. Verbs	.86	.46	-.35	.26	.75	.83
4. Modifiers	.72	-.06	.67	.52	.07	.54
5. Prepositions and Conjunctions	.54	.23	.40	.62	.28	.54
6. Noun Determiners	.33	.31	.54	.60	.61	.60

*Minus sign indicates significance in favor of control group.

r.05 = .19 r.01 = .25

expected direction. Exceptions to this occur where negative correlations are found between reading test scores on the pre-test and a negative correlation for the group doing modifier deletions on the tenth exercise, and a significant correlation for the group doing verb deletions on the twenty-second exercise. The correlations for the tenth exercise appeared to be generally somewhat lower than those for either the first or the twenty-second. Correlations with the post-test which do not fit the pattern occur for the tenth exercise, first for the random deletion group and second for the modifier deletion group. In general it appears safe to conclude that the reading ability of ninth grade pupils is related to their ability to complete correct closures.

Table XVIII gives the correlation of reading test scores with the number of correct comprehension exercises for Exercises 1, 10 and 22 by deletion type. This table gives an indication of greater

Table XVIII

CORRELATION OF READING TEST SCORES WITH NUMBER OF COMPREHENSION ITEMS CORRECT FOR EXERCISES 1, 10 AND 22 BY DELETION TYPE*

Deletion Type	Pretest DAT VR			Post-Test Iowa Reading		
	1	10	22	1	10	22
0. No Deletion	.34	.05	.38	.39	.55	.50
1. Random	.01	.06	.41	.23	-.03	.75
2. Nouns	.59	.86	.70	.64	.60	.73
3. Verbs	.73	-.85	-.15	.64	.63	.79
4. Modifiers	-.22	-.04	-.04	-.05	-.23	-.68
5. Propositions and Conjunctions	.18	.15	.39	.43	-.11	.44
6. Noun Determiners	.59	.27	.19	.48	.41	.66

*Minus sign indicates significance in favor of control group.

r.05 = .19   r.01 = .25

variability than that for number of correct closures. Here only the noun deletion group produces all significant correlations in the expected direction, and those correlations are generally higher than those of the control group. The modifier group produced correlations all in the negative direction indicating that the effect of modifier deletions is such that the comprehension of the material in question is negatively affected by the modifier deletion. This may be seen as a function of, first, the relative difficulty of modifier deletions; and second, the probability that information given through modifiers may be more important than that of other parts of speech in answering comprehension questions.

DISCUSSION

The non-overt reinforced cloze procedure does not appear to affect the reading comprehension of students to a different degree than does non-reinforced reading comprehension exercises. This suggests that the value of the cloze procedure exercise for teaching reading comprehension is a function of the cloze procedure plus the method by which it has been delivered. We cannot expect the cloze procedure to produce gains in reading comprehension under all conditions of presentation. There are, however, striking differences between the present study and the study of Schneyer[31] which found no significant increases, and the earlier study of Bloomer[1] and that of Roosnick.[34] Subjects in the Schneyer study were reinforced by the teacher, who returned the corrected cloze fill-ins to the students on the day following the cloze procedure exercise. Scoring was done in terms of the precise word only, meaning that to achieve a correct score the student must fill in the exact word which the author intended. In the earlier Bloomer study, the students graded their papers themselves for the exact word, and then the papers were turned in to be graded by the instructor for correct meaning. This difference in procedure may in and of itself be of significance for comprehension gains, in that a pupil did not receive reinforcement under the Schneyer technique where a word was perfectly appropriate. Hence the student would be likely to generate some considerable frustration related to correct language usage. In the Bloomer study any frustration generated by using correct, but not the author's, language would be dispelled at the subsequent meeting by the regrading of the instructor.

In the Schneyer study as in the present study, the students were bound to complete all of the exercises; hence there was no benefit for an extremely accurate performance. In the earlier Bloomer study, however, pupils advanced through grade levels more rapidly if their responses reached a 95% accuracy level. This meant essentially that any frustration developed as a function of being involved with the cloze procedure could be ameliorated most effectively by good performance. It appears probable, therefore, that the cloze procedure-methodological combination which is effective will include two elements in addition to cloze material: first, reinforcement by the teacher for the correct response, regardless of whether the response is the exact word of the author; and second, a motivational scheme with which the pupil's movement through the exercises will be determined by the quality of his response.

SUMMARY

The present analysis suggests that the use of non-reinforced cloze procedure is not effective for producing an increase in reading comprehension. The cloze procedure as an effective teaching device therefore represents a combination of the use of the material itself in addition to methodology which would appear to contain at least two variables; first is reinforcement for a correct response regardless of whether or not it is the precise word used by the author. Second, a motivational arrangement whereby the individual's progress through the cloze procedure material is related to the quality of his responses.

It appears from the correlational data that the ability to make closures is related to the reading ability of the student. However, the ability to answer comprehension questions subsequent to a cloze procedure exercise is dependent upon the type of cloze procedure exercise. Cloze procedure deletions, such as noun deletions, apparently offer more information and are more like straight reading material than correlations of modifiers. On the other hand, the deletion of modifiers appears to distort the meaning of the passage to a sufficient degree as to make the quality of the reading skill of the subject irrelevant to the number of comprehension questions answered correctly.

CHAPTER IX

THE RELATION OF NON-REINFORCED CLOZE PROCEDURE TO THE INTELLIGENCE OF THE RESPONDENT

Most reading and verbal activities bear some relationship to the intelligence of the individual making responses. One hundred and six ninth grade pupils who were involved in the experiment also took as a part of the experiment the California Test of Mental Maturity. Scores on the CTMM could then be correlated with the measures of cloze procedure responding. For the present study the total number of correct comprehension questions was correlated with CTMM scores as well as the total number of correct closures in the training exercises. For sequential comparison purposes, number of correct closures on the first, tenth and twenty-second exercise for each individual was correlated with CTMM scores. To test final results, the CTMM was correlated with the number of correct questions on the recall final, the number of correct recall questions relating to exercises No. 1, 10, and 22, and the number of correct closures by syntax as well as by training deletion type.

The experimental function of the CTMM was to determine differences between the randomly assigned experimental groups. Table XXIX gives the mean language, non-language, and total IQ mean scores for each experimental group and the control group. There are no significant differences between the groups in intelligence.

For interpretation purposes high correlations between intelligence

Table XXIX

CALIFORNIA TEST OF MENTAL MATURITY

MEAN SCORES BY DELETION TYPE

		Test		
		16	17	18
Deletion Type	Statistic	Language	Non-Language	Total
0. No Deletion	$\bar{x}$	109.53	118.74	114.32
	SD	17.36	16.23	14.40
1. Random	$\bar{x}$	114.60	119.22	117.22
	SD	14.52	17.11	12.77
2. Noun	$\bar{x}$	111.71	122.71	116.00
	SD	17.03	22.72	16.85
3. Verb	$\bar{x}$	103.00	111.50	107.50
	SD	14.08	27.50	20.41
4. Modifier	$\bar{x}$	114.92	129.42	122.41
	SD	13.59	17.11	11.58
5. Preposition and Conjunction	$\bar{x}$	115.56	121.44	118.94
	SD	9.80	16.10	10.99
6. Noun Determiner	$\bar{x}$	109.07	119.20	114.40
	SD	15.82	13.06	12.63

Table XXX

MEAN CORRELATIONS FOR CALIFORNIA TEST OF MENTAL MATURITY SCORES WITH NUMBER OF COMPREHENSION QUESTIONS CORRECT BY TYPE OF SYNTACTIC DELETION

	Training Deletion Type					
CTMM Score	0	1	2	4	5	6
Language	.26**	.30**	.83**	.25*	.27**	.26**
Non-Language	.24**	.22*	.77**	.34**	.07	.29**
Total	.30**	.26**	.82**	.49**	.08	.32**
N =	23	16	16	13	19	19

* Significant beyond .05 level of confidence

**Significant beyond .01 level of confidence and various cloze scores indicate the subject, because of the nature of the deletion, uses his intelligence for solutions to the problems presented in the materials. Low correlations indicate the nature of the material is such as to preclude the use of intelligence for solution. Such a case might occur when the material is completely unintelligible or where individuals of low intelligence can do as well as those of high intelligence, as in extremely easy material or material where the language patterns are firmly established in all individuals.

Intelligence and Comprehension. Table XXX gives the correlations for the California Test of Mental Maturity scores with the number of comprehension questions answered correctly by syntax deletion. The majority of the correlations are significant. However, certain variations can be seen between materials of differing deletion types. It should be noted that the small number of cases for Type 3 deletions taking the CTMM precluded the use of this particular deletion type in this comparison. The data indicate that the deletion type appears to be a variable in the relationship between comprehension of a passage and intelligence. Type 2 (noun deletions) correlates very strongly, indicating that when nouns are deleted from a passage the understanding of the passage is largely dependent upon the intelligence of the individual. Those individuals who are more intelligent will tend to understand it better, those who are less intelligent will not. On the other hand, if we consider Type 0 as a base line in the relationship between intelligence and reading comprehension, random deletions or the deletions of noun determiners do not significantly affect the relationship between reading comprehension and intelligence. This would seem to indicate that the disruption in meaning here is relatively insignificant and for ninth grade pupils random deletions at the ten percent level or deletions of noun determiners do not have a significant impact upon the intelligibility or the comprehension of the paragraph for the full range of individuals.

However, when we examine Type 5 deletions (prepositions and conjunctions) we observe a somewhat different phenomenon. Here the correlations all tend to be low, two of them are not significant. This would suggest that another variable is operating such that intelligence no longer enters into the process of comprehension through these materials.

One hypothesis might be that the materials have been rendered unintelligible to any degree of intelligence. However, since the degree of comprehension questions answered by individuals who have been trained with Type 5 deletions (prepositions and conjunctions) is markedly higher than most of the other situations, and since this is a function word and contributes relatively little to the meaning, this hypothesis is not tenable.

One might also hypothesize that the deletion of a preposition or conjunction might involve a certain interruptive process for the more intelligent individuals, but on the other hand would not affect the comprehension of the less intelligent individual, perhaps through his lack of recognition that something had been left out. Here again the number of correct comprehension questions for Type 5 is relatively higher than for other groups, indicating there is relatively little disruption on the part of the more intelligent individuals.

Evidence points to the fact that the deletions of prepositions and conjunctions tends to make the material somewhat easier to comprehend and it may be that material which contains prepositions and conjunctions, as the Yoakum Readability Formula would suggest, may tend to inhibit the reading of individuals of lower IQ and ostensibly also lower reading ability. The elimination of these therefore might be seen to serve as clarification for the individuals, the subscript line acting in many senses as a rather pronounced punctuation mark.

Further, in explaining this phenomenon we must also consider the fact of the focusing effect of the cloze procedure. It might appear as a function of the present data that where the cloze procedure forces the individual to focus on material, but at the same time does not offer him a challenge in terms of extreme loss of meaning, that this would be beneficial to individuals of lower IQ and tend to make their performance somewhat more consistent with that of individuals of the upper IQ, who may be seen as already approaching the upper limits allowed by the material. Thus where the deletion is prepositions and conjunctions any relationship between the comprehension of the material and the intelligence of the individual reader tends to be irrelevant.

Intelligence and the Sequence of Exercises. The correlation between CTMM scores and the number of correct comprehension questions for three of the sequence of twenty-four exercises was selected. These are Exercises No. 1, No. 10, and No. 22. As has been explained above in Procedure, Exercise No. 1 is not a single exercise but may have been any one of the twenty-four. It is simply the first exercise which was done by the individual. Similarly, exercise No. 10 is the tenth in the sequence, and exercise No. 22 the twenty-second in the sequence, although they may have been any of the twenty-four exercises presented. Table XXXI gives the correlations between the CTMM and the number of correct comprehension questions for the first, tenth and twenty-second exercises. Here no particular trend is noted, and the data for all of the deletion types is combined. All of the correlation coefficients are significant. Those for the language sections of the test are approximately equal to those for the non-language sections of the test, as well as for those of the total of the test. These data indicate that we are getting no gross over-all sequential effect of the sustained non-reinforced cloze procedure and intelligence. One would surmise therefore that consistent use of non-reinforced cloze procedure materials with a group of individuals does not disproportionately affect any subjects at different levels of intelligence.

Table XXXI

CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY

WITH NUMBER OF CORRECT COMPREHENSION QUESTIONS

FOR EXERCISES 1, 10 AND 22

	Exercise		
CTMM Score	1	10	22
Language	.45	.23	.38
Non-Language	.25	.30	.44
Total	.38	.25	.60

Intelligence and Recall of Comprehension Items. One would typically expect some relationship between recall of items and the intelligence of the individual learner. For the present study the recall final was given one and one-half weeks subsequent to the conclusion of the exercise section of the experiment. The recall test consisted of six of the twelve comprehension questions for each of the twenty-four exercises taken by an individual.

The correlations in Table XXXII indicate in general a moderately significant relationship exists between the intelligence of an individual as measured by the CTMM and the recall of items at periods from one and one-half to thirteen and one-half weeks later. Again these correlations seem to be most strong for individuals who had been trained on Type 2 (noun deletions). The level of relationship for those individuals trained on control Type 0 (no deletions) was approximately the same as the original correlations. Other correlations tended to be considerably lower, however. Two interesting differences did occur. The recall of items for Type 6 (noun determiners) tended to be less closely related to intelligence than were the immediate recall comprehension questions.

The correlations for delayed recall tend to be somewhat lower than those relating immediate recall and intelligence and suggest the possibility that recall is not as closely related to intelligence as immediate performance.

Relations of the CTMM with the number of correct comprehension questions on the recall final for the first, tenth, and twenty-second

Table XXXII

MEAN CORRELATION OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF COMPREHENSION QUESTIONS CORRECT ON RECALL FINAL BY TYPE OF SYNTACTIC DELETION

CCTM Score	Training Deletion Type					
	0	1	2	4	5	6
Language	.24*	.25*	.43**	-.08	.34**	.18
Non-Language	.25*	.09	.31**	.32**	.13	.13
Total	.30**	.23*	.42**	.23*	.25*	.19

Table XXXIII

CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CORRECT COMPREHENSION QUESTIONS ON RECALL FINAL FOR EXERCISES 1, 10 AND 22

CTMM Score	Exercises		
	1	10	22
Language	.47**	.21*	.24*
Non-Language	.03	.39*	-.09
Total	.20*	.41**	.13

exercises taken by the individual tend to support the notion, the reduced dependence of delayed recall and intelligence. More than half the correlations for the recall for exercises 1, 10 and 22 are not significantly related to the intelligence of the individual. This lack of significance appears to indicate that while intelligence may play some part in the delayed recall of items, this relationship is not particularly reliable.

Thus we find that the comprehension of items with varying syntactic deletions indicates a relatively high relationship between comprehension and recall of materials with noun deletions. The relationship is lower when function words are deleted, particularly for the immediate recall of material delivered in a context of missing prepositions and conjunctions. Further, during a long series of cloze materials, there appears to be no change in the relationship between intelligence and the ability to answer comprehension questions.

The delayed recall of items relating to previous exercises, with the exception of materials of noun deletions, does not indicate as strong a relationship between intelligence and the number of correct recall items as was found for the number of immediate recall items, suggesting that the relationship between intelligence and delayed recall is somewhat less than the relationship between intelligence and reading comprehension.

The Relation of Intelligence to Correct Closures. Table XXXIV indicates that the correlations between CTMM scores and the number of correct closures in training exercises is much the same as with comprehension questions. The correlation between correct random or noun closures and intelligence is relatively higher than for other closure types and the correlation with the number of modifier or preposition and conjunction closures is relatively lower.

Intelligence and Correct Closures on the Final Cloze Test by Type of Training Deletion. The final cloze test is a random cloze test. It is interesting to note that individuals who were trained

Table XXXIV
MEAN CORRELATIONS FOR CALIFORNIA TEST OF MENTAL MATURITY WITH NUMBER OF CORRECT CLOSURES IN TRAINING EXERCISES BY TYPE OF SYNTAX

	Training Deletion Type					
CTMM Score	0	1	2	4	5	6
Language		.48**	.24*	.08	.38*	.06
Non-Language		.31**	.42*	.20*	.11	.56*
Total		.31**	.60*	.21*	.25*	.41*

Table XXXV
MEAN CORRELATION FOR CALIFORNIA TEST OF MENTAL MATURITY WITH TOTAL CORRECT CLOSURES ON THE FINAL CLOZE TEST BY TYPE OF SYNTACTIC DELETION

	Training Deletion Type					
CTMM Score	0	1	2	4	5	6
Language	.37**	.09	.26**	-.05	.36**	.24*
Non-Language	.19	.13	.25*	.38**	.25*	.19
Total	.34*	.17	.40**	.27**	.33**	.25*

with random cloze tests have extremely low correlations between intelligence and the number of correct responses made in the final cloze test (see Table XXXV). During training, however, this relationship was not in evidence as significant correlations were obtained between training in random closures and intelligence. The indication here is that this is a training effect, in that these individuals are being tested in the mode in which they were trained and hence as the training has been effective the requirement for intelligent behavior is somewhat lessened. Individuals who were trained on specific syntax deletions or no deletions at all generally tended to indicate moderate significant correlation between intelligence and the ability to perform on the final cloze test.

Intelligence and the Number of Correct Closures for the First, Tenth and Twenty-Second Exercises in the Training Session. If the hypothesis that exposure to non-reinforced cloze procedure materials in fact trains an individual to perform closures, one would expect a certain dimunition of correlation coefficients between intelligence and the number of correct closures between the first and the twenty-second training session. This is evident from the data presented in Table XXXVI. The differences, however, are not significant, but since here we are dealing not only with individuals trained via the random cloze material but also individuals trained by other syntactic deletions, one would expect any differences to be occluded by this particular fact. This conjecture is upheld by the correlations for Type 1 (random deletions) and the language section of the CTMM. For the first exercise the correlation is .51; for the tenth exercise it is .45; and for the twenty-second it is .33, showing a gradual diminution in the relationship. The .33 with twenty-six degrees of freedom is not significant at the .05 level.

The proposition that individuals trained on random cloze procedure learn to make closures more effectively and hence are not required to use as much intelligence on the final cloze test as individuals trained through other procedure is not substantiated when one looks at non-language intelligence, or a total which reflects both language and non-language. The relationship between non-language intelligence and the number of correct closures moves from -.20 in the initial exercise to .30 in the final exercise. Neither of these are significant. However, the contributory effect of non-language intelligence shifts the total scores from .24 for the relationship between the CTMM total and the number of correct closures on the first exercise to a correlation of .50. On the twenty-second, this last is of course significant at the .01 level of confidence. This particular trend for the non-language and the total CTMM scores is not followed by the combined total groups.

Table XXXVI

CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY

WITH NUMBER OF CORRECT CLOSURES

FOR EXERCISES 1, 10 AND 22

CTMM Score	Exercises		
	1	10	22
Language	.40**	.20*	.31**
Non-Language	.38**	.42**	.21*
Total	.57**	.39**	.39**

Intelligence and the Number of Correct Closures on the Final Cloze Test by Syntactic Type. This comparison indicates the effects of the syntax of the deletion on the final cloze test itself, regardless of the syntax of the training procedure. In general, significant correlations are found between intelligence and the number of meaning-carrying words correctly fill in (see Table XXXVII). However, when we consider function words, particularly prepositions and conjunction, the relationship does not hold. This would suggest that function words or grammatical structure are basically learned by every individual and hence do not represent an intellectual type of problem.

Table XXXVII

CORRELATIONS OF CALIFORNIA TEST OF MENTAL MATURITY

WITH NUMBER OF CLOSURES BY SYNTACTIC TYPE

ON CLOZE FINAL TEST

	Type of Deletion				
CTMM SCORE	2	3	4	5	6
Language	.41**	.39**	.20	-.15	.18
Non-Language	.37**	.15	.26**	.11	.27**
Total	.47**	.40**	.29**	.01	.25*

r.05 = .20
r.01 = .26

SUMMARY

The CTMM was given to 106 of the population in the experiment. No significant differences between the groups were found. In general correlations of cloze scores and intelligence were positive and significant. However, in dealing with extremely difficult or extremely simple closures the correlations were low and in some cases negative. This phenomenon with the more difficult materials suggests that the deletions tend to decrease intelligibility of the selection to the point that intelligent behavior represents no advantage. Where the low correlations occurred with extremely simple materials, it appears that the language patterns are either well learned by all individuals or the deletions have the effect of shortening sentences to reduce readability.

CHAPTER X

AREAS FOR FUTURE EXPLORATION OF CLOZE PROCEDURE

Study of the cloze procedure and its utility is relatively new and hence has been confined to the areas of readability, comprehension testing and reading comprehension improvement. The study by Hafner has suggested a number of other potential uses for cloze. The present chapter proposes to suggest several additional uses or improvements of present uses which might be fruitful areas for exploration.

Cloze Readability and Comprehension Testing. The value of the cloze procedure as a readability predictor has been well established by Taylor. Reviews of cloze research such as that of Rankin show an increasing range of possibilities for experiment in cloze procedure readability. The present cloze techniques allow for putting materials in rank order for any particular group of individuals. It does seem surprising that some attempt has not been made to generate a standardized scale so that readability by the cloze procedure might be converted to some useful term such as grade level. Part of the reason may be that grade level is perhaps most useful at the very early grades and we find that cloze procedure is of relatively little use at early grade levels. Present first grade materials with many repetitions is much more difficult than second grade material, at least for adults. However, even in terms of grading material at the higher grade levels one must first of all establish a standard for the cloze procedure itself. The standard must be in terms of a percent of deletions as well as the types of deletions. One must also make the decision as to whether to be concerned with the precise word written by the author or with any correct meaning the respondent may supply. As Taylor suggests, these two scores correlate extremely well with each other. However, certain applications are suited to generalized meanings whereas others, such as the measurement of readability, may be more suited to specific meanings in that the scoring becomes considerably easier.

The scaling system in turn would be somewhat complex in that it also necessarily should take into account the reading level of the individuals used to grade the material, or use large sample techniques.

One problem arises with the use of cloze procedure in the classroom as a readability measure. If one takes a present class to establish the readability of material, we know only ex post facto that the material is either too difficult or too easy. These problems notwithstanding, a scaling system would be of great value because of the high reliability with which cloze can be scored, and the relative flexibility of cloze to adapt to the evaluation of many kinds of situations.

As with the application of any readability formula, more experimentation is required to determine the proper reading-teaching level. The question remains: "Shall we use relatively easier material and make the child feel comfortable in reading, or shall we use relatively more difficult material to make the child feel challenged by what he is doing." Probabilities are that different applications apply to different children or different types of children. The appropriate rules for this kind of application are still missing. There clearly must be some systematic relationship between the child, the situation in which he is reading, the nature of the material he is reading, and the reading level or difficulty of the material, and his teacher.

Cloze Procedure Comprehension Testing. The second area well established in cloze research is that of the testing of reading comprehension. Cloze provides reliable mechanism for testing comprehension, and here again there is a requirement for standardization. Cloze as a testing device can be readily administered and scored by teachers at a moment's notice, and yet to be in practical terms for a teacher a series of paragraphs, scaled for the number of errors by children of varying reading capacities, would be in order. Standardized decisions relating to type of syntactical deletion or percent of deletion would be a function of the author per se, but certainly such a device should prove extremely useful.

Cloze Procedure as an Intelligence Test. The study has shown a relationship between the number of correct closures and the intelligence of the responder. This relationship has been particularly strong where nouns were deleted from the material. Given the premise that intelligence is related to verbal ability, and verbal ability in turn related to the frequence of the vocabulary an individual uses, the correction factor for the frequency of occurrence of the word substituted might well serve to increase the correlation and hence the efficiency of the cloze as a measure of intelligence. Such a test would be related in part to an individual's ability to read. In addition, such a test would not necessarily be culture-free. The scores would depend, as would all cloze procedure responses, upon the language patterns of the individual. On the other hand, the test would be simple to administer, and would in turn be easily corrected for the individual's reading level. It would give an additional bonus to the individual who tends to use language in a creative fashion as the tendency to use clichés would mitigate against high scores. The problem with using cloze, as with nearly any other intelligence test, is that it would also be possible to train an individual to achieve a better score.

The use of the cloze procedure as a device to measure intelligence is predicated on the fact that one should be able to increase the relationship between other measures of intelligence and the cloze procedure by weighting responses according to their frequency of occurrence. It would appear similarly that cloze procedure with the deletion of function words such as prepositions and conjunctions, or noun determiners, would probably be ineffective.

Cloze Procedure as an Aptitude Test. Based upon the premise that an individual's language and ability to handle concepts is superior in an area of interest and potential specific aptitude tests using cloze procedure could be devised, these aptitude tests would afford, as did the intelligence test, a weighted score for infrequent or specialty bound words. Materials of varying levels in the social sciences, the sciences or any vocational area, can be graded in difficulty to give an indication of an individuals ability to handle conceptualizations in any particular field. A well-developed battery of these materials with several levels of materials in important areas, with corrections for the individual's age and reading ability, might well afford an effective predictor of specific academic or vocational success.

Cloze Procedure as a Test of Syntactic Ability. While individual syntax is no longer considered a serious problem, where an individual syntax is to be converted to standard syntax it would seem that the cloze procedure could be used as a test of standard syntactic ability, in that certain syntactic problems could be incorporated in the blanks of one or several cloze procedure passages. Reading level and syntactic ability would tend to interact in such an instrument. However, this interaction could be largely avoided, except for cases of extreme reading disability, by making the cloze material selected of extremely low readability. Since cloze is of little value below fourth or fifth grade level because of the language pattern development of children, material whose initial reading level was of second or third grade level before cloze mutilation should present relatively few reading problems except to the most seriously disabled readers. Clearly, as the level of the child to be tested increases the readability of the material may also be increased but in any case should remain several grade levels below the mean for the children to be tested.

It would be possible to generate a wider range test from a single selection. Syntactic ability under these circumstances would not be independent of an individual's reading ability and therefore sub-scores relating to one's ability to fill in words of particular syntax or to solve particular syntactical problems would be related to the reading ability of the child. For instance, the child at fourth grade reading level who is able to solve seven out of twenty auxiliary verb problems might be operating at normal level, whereas the child at seventh grade level might be expected to resolve correctly fifteen of the twenty auxiliary verb problems. Scoring would necessarily have to be in terms of the syntax of the supplied word as well as in terms of the meaning of the supplied word. The syntactic problems would have to be chosen with great care to fit most probable errors; however, it would appear that this kind of test would afford the testor the possibility of drawing both reading level and syntactical level from the same instrument by using a context which, for an individual, is relatively more normal.

The present experiment has indicated some potential problems which may result in lowered reliability of cloze materials either as a readability testing device or as a reading comprehension device where different passages are used to measure the same group of children. The first circumstance occurs where the style of the author is relatively even, such that the sentence lengths are approximately the same, and the structure of the sentences is also relatively the same. Where every nth word, the deleted word, represents the length of the sentence, then each deletion will be of the same syntax, and a distorted result would be obtained. The second circumstance is similar, but should by chance an inordinately large block of deletions be of the same syntax or syntaxes of similar difficulty, one might also get a disproportionate result. It is clear that the danger of this difficulty increases as the proportion of closures per sample size decreases. Hence in a large number of situations, one deletion in five words would tend to be somewhat more reliable than one deletion in ten words, in turn more reliable than one deletion in twenty words, etc. Users of cloze material must guard against this problem. The relative proportion of deletions of any particular syntax should approximate the proportion of the occurrences of that syntax in the material. Such a check should be made on any random cloze material. A similar check is not necessary with syntactic deletions as long as appropriate compensations are made for the relative difficulty of the syntax deleted.

The Cloze Procedure as a Controlled Projective Device. The cloze procedure as it stands bears a relationship to sentence-completion type personality tests. Cloze procedure responses appear to vary with the personality of the individual filling in the closures, and seem to form a consistent pattern. Sentence completion techniques bring up a consistently varied sequence of problem situations through the cue stem. The cloze procedure, however, could be used to measure the consistency of a particular reaction by giving a longer contextual association relating to a particular kind of problem. The cloze context would determine the nature of the response to some degree by determining the theme of the problem. Problems would be built in such a way that an individual may make a number of possible answers depending upon his basic hostility, insecurity or security, or sociability, or other dominant personality theme. The examiner would be concerned not only with the nature of the response but also with the consistence of the response to a particular cloze passage or through several different cloze passages.

This would serve to give us a much more reliable estimate than is presently found in sentence completion type materials. The following represent examples of materials:

His ____________ did not like him and others thought
he was ___________. When he went to ______________ he was
always _____________. Sometimes __________ would ask him
to ___________.

or:

Suzan was a ____________ little girl. Her mother and
father often asked her to ___________. Her mother and
father were never __________. She liked to play best
with ___________ and sometimes she ______________.

Further, it would be possible to develop an objective scoring device using the cloze procedure. One method of developing an objective scoring system would be to use the semantic differential to estimate the position of various words on a continuum. Each word substituted in the cloze passage might therefore be graded on a number of dimensions. The relative sum of words for each of the dimensions, the sum of the scores for any particular dimension, would give the propensity of an individual to move into or to be affected by that particular dimension. Standard scores could be developed for individuals of varying age levels.

The selection of the type of response to be omitted should be fairly carefully done in using the cloze as projective. Hints as to the form or the particular thing desired should be deleted. Hence if a single noun is deleted and it reoccurs it should be deleted each time. It would be most common to delete verbs or modifiers as these tend to convey the greatest amount of emotion. The deletion of structure words would allow the reader to determine the direction that the material would take.

For the purpose of the projective cloze, the material must be of emotional significance, dealing with a swlected number of problems which represent significant dimensions of personality. The material should be designed such that stories can evolve in different ways if different words are selected. Thus the character of the story will be a function of the individual producing the story.

Cloze as a Teaching Instrument. The cloze procedure at this juncture appears to have four possibilities as a teaching instrument. In each case the cloze procedure would be handled somewhat differently to produce desired results. The four possibilities are: the teaching of grammar syntax; the teaching of composition; the increase of reading comprehension; and the teaching of content.

Grammar and Syntax. The data in the present study with non-reinforced cloze procedure indicates an increase in the ability to make closures of certain specified syntaxes of the syntactic categories used. Increases were noted in the ability to fill in prepositions and conjunctions, noun determiners, modifiers, and in some instances, verbs. Random cloze procedure training of a non-reinforced variety also produced an increase in the ability to fill in words of several syntaxes. It appears as though the non-reinforced cloze procedure, however, is ineffective with nouns, and relatively ineffective with verbs.

It should be noted that cloze procedure exercises bear a certain similarity to the sentence type exercise which is customarily used for the teaching of grammatical structure in workbook-type materials. Workbook materials typically consist of a sentence in which a blank has been left for the individual to fill in, involving some syntactic problem. Sometimes the blanks are presented with multiple choices and sometimes with constructed responses. The probability of a correct response is controlled by the selection of the technique of presenting the correct answer. Further, these materials are typically reinforced by an answer key at the back, or by teacher or student correction of the material subsequent to the working of a page. These materials have the disadvantage that they are discontinuous sentences while normal speech or writing is more continuous in form. This objection could be avoided through the use of the cloze procedure.

Use of the cloze procedure has not only increased the number of closures of the appropriate syntaxes but also the number of specific words the author used. This suggests that the cloze procedure tends to produce both more conventional language patterns and a higher degree of syntactic accuracy.

It should be also noted that the use of non-reinforced cloze procedure is relatively inefficient. It appears necessary to attempt an experiment using cloze procedure for the teaching of syntax using a reinforced form. Theoretically the use of reinforcement would be more effective.

The use of the cloze procedure as a technique for teaching syntax requires diagnostic technique. Within any particular group many syntactic errors would be subcultural in that they would be held in common by a particular group of students. It is essential to determine beforehand which errors are common for a particular child and which are common to a group of children. A teacher might devise a test again using cloze procedure where the teacher specifically deletes words which relate to a syspected syntactic error. Needless to say there should necessarily be several instances of any suspected syntactic error included in the testing materials. The deletions need not necessarily be limited to a single word. Phrases of two or three words where these involve certain syntactic and grammatic problems might be deleted. It should be cautioned that multiple word deletions be limited to upper grade students who have some experience with the language and are able to substitute longer phrases with some degree of accuracy. Once the test has been used to determine the problems of a syntactic nature for any particular child or group, cloze procedure materials can be developed to correct a particular set of substandard syntaxes. Eighteen to twenty such cloze exercises of about fifty responses each concentrating on specific syntactic problems, reinforced by a list of correct responses and also reinforced by subsequent checking for correct meanings on the part of the teacher, would be an effective method of teaching standard syntax to a child. The correct answers for the child would be any words which are meaningful which also involve the appropriate syntax, or in turn words which would adjust or change the syntax in an acceptable manner. An acceptable answer must be in keeping with the total paragraph and not simply restricted to the particular sentence with the blank in it. One real advantage of this use of cloze procedure is that the child must observe not only the syntax within the specific sentence but also the context in which the sentence occurs.

Such a procedure would essentially involve experimentation with the cloze form and other potential techniques for teaching standard syntax before widespread adoption of such a technique would be advisable.

__Composition__. The teaching of composition is, from the present point of view, almost an opposite kind of problem from the teaching of syntax. Syntax involves the use of regular and relatively conventionalized forms, whereas the goal in composition is the use of unusual and inventive ways of expressing oneself. Again, contrary to the problem of syntax, the syntactic forms which carry the greatest weight in the composition are such words as verbs, adjectives and nouns. The variety and unusualness of these words tend to be related to the polity of the composition.

The English teacher who is responsible for the teaching of composition is beset with the problem that his students' work does not match the standards which he desires. A large bulk of the literature which relates to composition expresses various techniques for increasing the quality of student work. The fact that this kind of literature is continuing is testimony to the relatively limited success of the techniques offered. The difficulties which the students appear to have lie in three general areas. The first is organization of materials. The students appear not to be able to get a generalized over-view of the material and hence the material they write is poorly organized. Second, English teachers bemoan the number of cliches, the paucity of the language of the students. The written expression of pupils appears to take a very conventional form and hence does not captivate the audience nor express in subtlety the limits of the subject. The third area for concern is the relatively poor syntax of the pupils. Some pupils seem to make systematic sets of grammatical errors and these appear to be very difficult to correct.

We can see a relationship between the teaching of syntax and the over-use of cliches. The pupil tends to learn to use his language in patterns and hence does not weigh each specific word which he utters or writes, but rather begins on a particular language chain which will lead him to his desired end point and simply fills in from his learned language pattern and does not think about choosing the words in between. We can see that many of the syntax and grammar teaching exercises we give to children do exactly this--that is to say, they teach the child an appropriate language pattern which is the same as a cliche, the difference being that for present purposes the syntactical language pattern is an acceptable and common convention whereas the cliche because of its over-use becomes qualitatively unacceptable. In essence, then, the language teacher in some measure works against himself by teaching grammatical form as a language pattern, in the same way that the teaching of idioms in a foreign language reduces the degree of freedom of the child in terms of choice for an expression of a particular concept. If we look at the complaint of the language teacher, that the child tends to use few words and to use cliches, we see that this complaint resolves itself to the fact that the words used by the child are in common patterns and are also words which are highly frequent and partly as a function of this, also low in emotional impact. If we can judge from the closures in the present experiment, the problem of learning a syntax is relatively minute and hence pupils' composition could be upgraded considerably by increasing the richness of their language. The solution therefore to part of the problem of composition is to design materials which will increase the variability and breadth of the child's language usage. Here again if we are to judge from the vocabulary studies, the number of words with which a child is familiar and can identify is quite large. The problem resolves itself to the fact he does not use these words.

Studies with word association techniques indicate that when an individual is confronted with words which have an emotional impact for him, a certain group of people, depending upon their particular personality, will tend to respond in kind; that is to say, if given a highly emotional stimulus, most people will respond with a highly emotional response; given a relatively neutral stimulus, most people will respond with a highly neutral response. If we can therefore develop materials which are compositional in nature which give emotional stimuli, we can expect an increase in the amount of emotional response we will get and theoretically if we give an individual infrequent words we will by the same token tend to get less frequent words in return.

It is clear that the number of external stimuli in a composition situation is relatively limited, usually merely a theme or topic is given to the individual and he in turn will respond to this by producing a relatively long chain of responses on or related to this particular theme. This presents a problem in composition in that the individual is allowed to move from the theme and the stimulation produced by it through his own series of response chains, which in turn will stimulate him to move toward his own medium and hence his compositions will have relatively smaller degree of variability in them.

Given the fact that an individual responds to verbal stimuli by responses in kind, and since we would wish to continue the production of highly emotional and infrequent responses on the part of the individual, in order to train the individual to more complex language patterns we will have to present not just a theme as a set of stimuli but also a series of continuous stimuli throughout the composition which will reinstigate the writer to more emotional and less frequent language. The obvious detriment to this kind of system, however, is that the individual will then become dependent upon successive sets of stimuli throughout his composition and will not be trained to produce his own material. However, as we withdraw the amount of stimulation in a verbal word association situation we find that the quality of responses remains essentially the same, although the variability or predictability of responses decreases. Thus, after we have trained an individual to respond to emotional stimuli by producing emotional words and respond to infrequent stimuli by producing infrequent words, we can gradually withdraw by a technique of fading similar to that used by Skinner, and make the individual more independent and more able to generate his own long sequence patterns.

The cloze procedure offers a technique for accomplishing just this end. Typically in the cloze procedure an individual draws from his own stock of words and his own language patterns to fill in the blanks which are left in the material. Typically he tends as well to match the style of the author, which accounts for the large number of correct responses which children in junior high and high school will make to cloze procedure material. We see therefore that the cloze procedure offers a technique for shaping the responses of the individual. By giving the pupil cloze procedure material in which the stimuli or the cloze cue (that is, the word immediately preceding the blank which is to be filled in) is a highly stimulating word, we will be likely to have the individual draw upon his own background of highly stimulating words to fill in the blank. If we train the individual at one level deletion to fill in words, and then we gradually decrease the amount of stimulation and increase the amount of independence which the individual has, we find that we can fade the cloze procedure from the composition situation and successfully train the individual to make compositions of more highly varied material. Further, the fading situation will also be of some assistance in teaching the child to organize material, because the first material he gets will be highly structured and organized. As he continues to respond, he will have to insert more of his own structure and organization. Hence in some measure the use of the fading cloze procedure as a training device will afford the child the use of richer language patterns and in some measure assist him in the organization of his thoughts. The cloze procedure also offers sufficient structure so that the individual makes a miniscule number of syntactical errors. In the present study the number of syntactical errors made in the cloze procedure materials is less than one percent. However, one would expect that syntactical errors would increase as the fading increases. Since each student's paper would be graded for syntax as well as for other variables, these errors would be detected early and the child could then profit by syntactic cloze training.

Reading Comprehension and the Cloze Procedure. The present study with the cloze procedure indicated no growth in reading comprehension on the part of the participants in the experiment. Other experiments cited before (Bloomer[1]) which handled the cloze procedure in a different manner found that cloze procedure appeared, at least at the college and the sixth grade levels, to be related to growth in reading comprehension. A recent study by Schneyer[34], however, found no particular effects from prolonged cloze procedure. The difference in these studies indicate that the manner in which the cloze procedure is presented will either produce or not produce effective results in increasing reading comprehension. Carefully prepared motivational conditions are required for the production of reading comprehension. The question as to whether results in increased reading comprehension under motivational conditions are a function of cloze procedure or merely a function of the motivational conditions as applied to any type of reading comprehension material remains to be answered.

The motivational components which appeared to be effective when used in conjunction with the cloze procedure are the following.

The first is the use of reinforced cloze procedure. This involves the use of a correct answer list used by the individual to score his material subsequent to completion of the cloze procedure exercise. The answer list gives the word which the author put into the particular place in the paragraph, and hence the student receives an immediate score by grading the paper himself, and furthermore obtains specific feedback in relation to the various blanks which he had questions about.

The second motivational factor is regrading for meaning by the instructor. This can be accomplished by having the instructor collect the papers at the termination of the grading by the student, and in turn, using a somewhat broader key which has not only the correct word written by the specific author but also other words of appropriate meaning. The instructor's grade is typically somewhat higher than that of the student himself, hence giving additional reinforcement when returned to the student. In this sense, the student first receives reinforcement for completion of the material in a manner meaningful to him; secondly, reinforcement for his approximation of the author's language; and third, reinforcement from the instructor when the papers graded for correct meaning are returned to him. It also appears that some form of chart to make comparative scores available to the student so that he can watch his progress through the cloze procedure material is of value, He can thus reinforce himself as he progresses in terms of the number of correct responses.

Third, the material should be graded in difficulty; hence the individual will start with relatively simple material and continue until he is reading and filling in material which is very close to his present reading level. The individual can therefore watch himself progressing in the grade level of his ability as he continues through the program.

Fourth, movement up the grade level should be contingent upon the number of correct responses made by the individual student. Basically then the material should be supplied to the student in such a manner that there are several bits of material in depth at any particular grade level. The criterion for movement should be relatively high, approaching 90 to 95% correct responses. The individual who fails to achieve such a score will therefore remain with material at the same grade level. The individual who achieves at the 90 or 95% correct response level will move to a level of more difficult material. Thus the individual student's rate through the program is contingent upon the quality of his responses. The individual who does very well is reinforced by a number of correct responses, and is also reinforced by his movement from grade level to grade level through the end of the program.

At this point the content of the material in terms of its interest to the individual appears to be irrelevant. More work would need to be done to establish whether or not the effectiveness of the material can be controlled by the interest level of the material.

There appears to be a further limitation in relation to the age group for which this material is effective. Cloze procedure will not work for children who cannot spell or write in the appropriate missing words, though an attempt at the use of oral cloze with younger children might prove to be interesting. Further limitation is that cloze procedure is ineffective with pupils who have not developed fairly strong language patterns. Younger children appear to repeat the cloze cue rather than to fill in a word which will make the material meaningful and hence do not develop systematic meaningfulness from cloze material. A mental age of about ten or higher is required to function adequately with cloze procedure material. Below this level the probability of making errors because of poorly developed ability to manipulate language is extremely high, resulting in frustration and producing no real benefits in ability to read. Therefore above a certain level children will learn to comprehend more effectually with properly designed cloze materials. Here again it must be cautioned that more experimentation needs to be done with the cloze procedure as a reading comprehension device.

Cloze Procedure and Content Learning. Experiments with cloze procedure indicate that cloze procedure is effective for increasing the immediate comprehension of materials. The present study indicates only mild relationships occur between comprehension increases and cloze procedure with the use of non-reinforced cloze procedure. The use of cloze procedure to increase understanding of content should be reinforced. Pilot studies show that extremely difficult cloze procedure material is ineffective in teaching content.

If we look at cloze procedure in relation to programed instruction, we find that reinforced cloze procedure has all the elements of good programed instruction. First, the cloze procedure involves content material and can be used to deliver content to an individual, a prime requisite of programed instruction.

Second, cloze procedure demands responses of individuals, which is the second criterion for programed instruction; and third, cloze procedure presents reinforcement for correct responses.

Some limitations in relation to the organization of the material made into cloze content material should be expressed. The material must move carefully and sequentially from lower order concepts to higher order concepts. It is often difficult to find written material of this nature.

The second criterion would be that the material be relatively redundant to produce a high order of correct responses on the part of the individual. This again tends to limit the amount of material available. However, material certainly could be easily written to meet these criteria.

Third, cloze deletions of teaching material content should involve the use of nouns, verbs and adjectives in greater proportion than the function words, because we are concerned with the individual's ability to discover and fathom the carriers of meaning rather than their syntactic relationship in particular situations. Occasionally, however, a significant word might fall in one of the function word categories. It should be noted that the selection of closures is not random but rather deletions are made on the basis, first, of their importance as carriers of meaning, and second, their presence in other situations or their value as solutions to conceptual problems.

The cloze procedure will probably be somewhat less effective than programed instruction of the Skinnerian type in producing verbal responses patterned directly upon the content of the material. On the other hand, cloze procedure can be manipulated such that the individual will be more likely to conceptualize the material.

For the purpose of increasing the probability of reinforcement a key word list (a scrambled list of all the correct answers) may be placed upon the page so the individual learner can use the list as an aid to achieving correct responses. Reinforcement with a key word list would be the specific word deleted, particularly where noun deletions are used. With deletions of adjectives and verbs, the instructor may use a second scoring key which would allow adequate substitutes.

As with other applications, the use of cloze procedure as a technique for increasing an individual's knowledge of content would require considerably more experimentation. Certain indications in the literature appear to suggest that cloze procedure will increase immediate comprehension, but there must also be a concern with the delayed recall of content material learned in a cloze procedure exercise. This appears to be an area of cloze procedure which is worthy of experimental effort.

SUMMARY

A large number of potential applications of cloze procedure, each requiring considerably more research, are evident.

First, standardization of cloze procedure as a readability measurement instrument and as a reading comprehension device would

make it a more useful tool for the classroom teacher.

Second, cloze can be adapted to a number of other testing situations--as an IQ test; a specific aptitude test for career advisement; a test of standard syntactic skill; and as a projective device, appear to have great promise.

Third, cloze as a teaching instrument appears to have potential for teaching standard syntax, composition, reading comprehension and content.

In each case exploration of an avenue of cloze application should be conducted in strict experimental form.

BIBLIOGRAPHY

1. Bloomer, R.H., "The Cloze Procedure as a Remedial Reading Exercise," Journal of Developmental Reading, 5:173-181, Spring 1962.

2. Bormuth, J.R., Cloze Tests as Measures of Readability. Unpublished doctoral dissertation, Indiana University, 1962.

3. Bormuth, J.R., "Cloze as a Measure of Readability," International Reading Association Converence Proceedings, New York: Scholastic Magazines, 8:131-134, 1963.

4. Bormuth, J.R., "Experimental Applications of Cloze Tests," International Reading Association Conference Proceedings, Newark, Delaware: International Reading Association, 9: 1964.

5. Bormuth, J.R. "Relationships Between Selected Language Variables and Comprehension Ability and Difficulty." Cooperative Research Project Number 2082, U.S. Dept. of Health, Education, and Welfare, 1964.

6. Brooks, S.T., Effects of Locus of Control and Anxiety on the Ability of Mentally Retarded Children to Use Context Clues in Reading. Unpublished doctoral dissertation, George Peabody College for Teachers, 1962.

7. Carroll, J.B., A.S., and C.P. Wilds, "An Investigation of Cloze Items in the Measurement of Achievement in Foreign Languages," A Report on Research Conducted under a Grant from the College Entrance Examination Board, Laboratory for Research in Instruction, Graduate School of Education, Harvard University, Cambridge, Massachusetts, 1959, 1-138.

8. Cofer, C.N., "Cloze Procedure in the Evaluation of Prose Stories Varying in Adjective-Verb Quotient," Technical Report No. 13 for Contract N.O.N.R. 595 (04), Department of Psychology, University of Maryland, College Park Md., 1957, 1-6.

9. Cofer, C.N., and P.M. Jenkins, "A Study of the Learning of Two Stories by Means of the Cloze Technique," Technical Report No. 14 for Contract N.O.N.R. 595(04), Department of Psychology, University of Maryland, College Park, Md., 1957, 1-6.

10. Coleman, E.B., "Improving Comprehension by Shortening Sentences." Journal of Applied Psychology, 46:131-134, April, 1962.

11. Coleman, E.B., and J.P. Blumenfeld, "Cloze Scores of Nomilizations and their Grammatical Transformations Using Active Verbs," Psychological Reports, 13:651-654, Dec. 1963.

12. Fillenbaum, S., and L.V. Jones. "An Application of Cloze Technique to the Study of Aphasic Speech," Journal of Abnormal and Social Psychology, 65:183-189. Sept. 1962.

13. Fillenbaum, S., L.V. Jones, and A. Rapoport, "The Predictability of Words and Their Grammatical Classes as a Function of Rate of Deletion from a Speech Transcript," Journal of Verbal Learning and Verbal Behavior, 1963, 2:186-194.

14. Fletcher, J.E., *A Study of the Relationships Between Ability to Use Context as an Aid in Reading and Other Verbal Abilities*. Unpublished doctoral dissertation, University of Washington, 1959.

15. Friedman, M., *The Use of the Cloze Procedure for Improving the Reading Comprehension of Foreign Students at the University of Florida*. Unpublished doctoral dissertation, The University of Florida, 1964.

16. Fries, C.C., *The Structure of English*, New York: Harcourt Brace, 1952.

17. Greene, F.P., *A Modified Cloze Procedure for Assessing Adult Reading Comprehension*. Unpublished doctoral dissertation, The University of Michigan, 1964.

18. Hafner, L.E., "Implications of Cloze." In Eric L. Thurston and L.E. Hafner (Eds), *Fourteenth Yearbook of the National Reading Conference*, Milwaukee: National Reading Conference, Inc., 1965, 151-158.

19. Hafner, L.E., "Relationships of Various Measures to the Cloze," In Eric L. Thurston and Lawrence E. Hafner (Eds.), *Thirteenth Yearbook of the National Reading Conference*, Milwaukee: National Reading Conference, Inc., 1963, 135-145.

20. Jenkins, J.J., "Communality of Associations as an Indicator of More General Patterns of Verbal Behavior," In T.A. Sebock (Ed.), *Style in Language*, New York: John Wiley and Sons, 1960, 307-329.

21. Jenkinson, M.S. *Selected Processes and Difficulties in Reading Comprehension*. Unpublished doctoral dissertation, University of Chicago, 1957.

22. MacGintie, W.H., "Contextual Constraint in English Prose Paragraphs," *Journal of Psychology*, 51:121-130, Jan. 1961.

23. Manis, M., and R.M. Dawes, "Cloze Scores as a Function of Attitude," *Psychological Reports*, 9:79-84, 1961.

24. Osgood, C.E., "Some Effects of Motivation on Style of Encoding," In T.A. Sebock (Ed.), *Style in Language*, New York: John Wiley and Sons, 1960, 293-306.

25. Rankin, E.F., Jr., *An Evaluation of the Cloze Procedure as a Technique for Measuring Reading Comprehension*. Unpublished doctoral dissertation, The University of Michigan, 1957.

26. Rankin, E.F., Jr., "The Cloze Procedure--Its Validity and Utility," In O.S. Causey and W. Eller (Eds.) *Eighth Yearbook of the National Reading Conference*, National Reading Conference, Inc., 8:131-144, 1959.

27. Rankin, E.F., Jr., "The Definition of Reading Comprehension," *First Yearbook of the North Central Reading Association*, University of Minnesota, Minneapolis: 15-31.

28. Rankin, E.F., Jr., "Reading Test Reliability and Validity as a Function of Introversion-Extroversion," *Journal of Developmental Reading*, 6:106-117, Winter, 1963.

29. Rankin, E.F. Jr., "Reading Test Performance of Introverts and Extroverts," In Ralph S. Staiger and Culbreth Y. Melton (Eds.), *Twelfth Yearbook of the National Reading Conference*, Milwaukee: National Reading Conference, Inc., 12:158-166, 1963.

30. Rankin, E.F., Jr., "Uses of the Cloze Procedure in the Reading Clinic," In J. Allen Figurel, (Ed.), *International Reading Association Conference Proceedings*, New York: Scholastic Magazines, 4:228-232, 1959.

31. Roossinck, P.L., *A Method for Improving Reading Comprehension*. M.A. thesis, The University of Michigan, 1962.

32. Ruddell, R.B., *An Investigation of the Effect of the Similarity of Oral and Written Patterns of Language Structure on Reading Comprehension*. Unpublished doctoral dissertation, Indiana University, 1963.

33. Salzinger, S., "A Study of Readability Measurement-Application of Cloze Procedure to Japanese Language," (English Abstract) *Japanese Journal of Psychology*, 28:67-73, 1957.

34. Schneyer, J. Wesley, "Use of the Cloze Procedure for Improving Reading Comprehension," *Reading Teacher*, 19 (Dec. 1965), 174-79.

35. Smith, D.E.P., *Grammatical and Logical Units for Increasing the Complexity of Paragraphs*, Unpublished manuscript, University of Michigan.

36. Smith, D.E.P., *A Theoretical Model of Reading Comprehension*, Unpublished Manuscript, University of Michigan.

37. Sukeyori, K., S. Portnoy, and R.S. Feldman, "The Effect of Order of Approximation to the Statistical Structure of the English Languagedon the Emission of Verbal Responses," *Journal of Experimental Psychology*, 64:52-57, July, 1962.

38. Taylor, W.L., "Cloze Procedure: A New Tool for Measuring Readibility," *Journalism Quarterly*, 1953, 30:414-438.

39. Taylor, W.L., *Application of "Cloze" and Entropy Measures to the Study of Contextual Constraint in Samples of Continuous Prose*. Unpublished doctoral dissertation, University of Illinois, 1954.

40. Taylor, W.L., "K.M. Readers Lend Hand to Science: Cloze Method Works in Written Korean and May Serve as a Tool for Korean Language Reform," *Korean Messenger*, 3:4-5, 1954.

41. Taylor, W.L., "Recent Developments in the Use of the Cloze Procedure," Journalism Quarterly, 33:42-48, 1956.

42. Taylor, W.L., "Cloze Readability Scores as Indices of Individual Differences in Comprehension and Aptitude," *Journal of Applied Psychology*, 41:19-26, February, 1957.

43. Weaver, W.W., An Examination of Some Differences in Oral and Written Language Using the Cloze Procedure. Unpublished doctoral dissertation, University of Georgia, 1961.

44. Weaver, W.W., "Theoretical Aspects of the Cloze Procedure," In Eric L. Thurston and Lawrence E. Hafner (Eds.), Fourteenth Yearbook of the National Reading Conference, Milwaukee: National Reading Conference, Inc., 1965, 113-132.

45. Weaver, W.W., "The Predictability of Omissions in Reading and Listening," In Emery P. Bliesmer and Ralph C. Staiger (Eds.), Eleventh Yearbook of the National Reading Conference, Milwaukee: The National Reading Conference, Inc., 1962.

46. Weaver, W.W., and A.J. Kingston, "A Factor Analysis of the Cloze Procedure and Other Measures of Reading and Language Ability," The Journal of Communication, 13:252-261, December, 1963.

47. Weinfeld, F.D., A Factor Analytic Approach to the Measurement of Differential Effects of Training: An Evaluation of Four Methods of Teaching English Composition. Unpublished doctoral dissertation, Harvard University, 1959.